Search for invisible Higgs-boson decays in events with vector-boson fusion signatures using 139 fb$^{-1}$ of proton-proton data recorded by the ATLAS experiment

The ATLAS Collaboration

DOI
10.1007/JHEP08(2022)104

Publication date
2022

Document Version
Final published version

Published in
Journal of High Energy Physics

License
CC BY

Citation for published version (APA):
https://doi.org/10.1007/JHEP08(2022)104
Search for invisible Higgs-boson decays in events with vector-boson fusion signatures using $139 \text{ fb}^{-1}$ of proton-proton data recorded by the ATLAS experiment

The ATLAS collaboration

E-mail: atlas.publications@cern.ch

Abstract: A direct search for Higgs bosons produced via vector-boson fusion and subsequently decaying into invisible particles is reported. The analysis uses $139 \text{ fb}^{-1}$ of $pp$ collision data at a centre-of-mass energy of $\sqrt{s} = 13 \text{ TeV}$ recorded by the ATLAS detector at the LHC. The observed numbers of events are found to be in agreement with the background expectation from Standard Model processes. For a scalar Higgs boson with a mass of $125 \text{ GeV}$ and a Standard Model production cross section, an observed upper limit of $0.145$ is placed on the branching fraction of its decay into invisible particles at 95% confidence level, with an expected limit of $0.103$. These results are interpreted in the context of models where the Higgs boson acts as a portal to dark matter, and limits are set on the scattering cross section of weakly interacting massive particles and nucleons. Invisible decays of additional scalar bosons with masses from $50 \text{ GeV}$ to $2 \text{ TeV}$ are also studied, and the derived upper limits on the cross section times branching fraction decrease with increasing mass from $1.0 \text{ pb}$ for a scalar boson mass of $50 \text{ GeV}$ to $0.1 \text{ pb}$ at a mass of $2 \text{ TeV}$.

Keywords: Hadron-Hadron Scattering, Higgs Physics, Vector Boson Production

ArXiv ePrint: 2202.07953
## Contents

1 Introduction ................................................. 1

2 ATLAS detector ........................................... 3

3 Data and simulation ......................................... 4

3.1 Data and triggers ........................................ 4

3.2 Event simulation ......................................... 4

3.2.1 $V + \text{jets and } VV + \text{jets}$ ....................... 5

3.2.2 $t\bar{t}$, $tW$, and single top ........................... 7

3.2.3 QCD multijet ........................................... 7

3.2.4 Signal ................................................... 7

4 Object and event selection ................................. 8

4.1 Object definitions and event cleaning ................. 8

4.2 Signature from vector-boson fusion ..................... 10

4.3 Event selection .......................................... 11

5 Background estimation ..................................... 12

5.1 $V + \text{jets background}$ ................................. 12

5.2 Multijet background ..................................... 17

6 Systematic uncertainties ................................... 23

6.1 Theoretical uncertainties ................................ 23

6.1.1 $W + \text{jets and } Z + \text{jets uncertainties}$ ....... 24

6.1.2 Signal uncertainties .................................. 25

6.2 Experimental uncertainties ............................. 26

7 Statistical model .......................................... 27

8 Validation .................................................. 29

9 Results ..................................................... 29

10 Conclusion ............................................... 38

The ATLAS collaboration ...................................... 49
1 Introduction

This paper presents a direct search for the decay of the Higgs boson [1, 2] into invisible particles. While the expectation for the branching fraction of invisible decays ($B_{\text{inv}}$) from the Standard Model (SM) is 0.12% due to the decay $H \rightarrow ZZ^* \rightarrow \nu \bar{\nu} \nu \bar{\nu}$, several scenarios of physics beyond the Standard Model (BSM) allow for anomalous and sizeable $O(10\%)$ values. In one class of models under consideration [3–15] the Higgs boson might decay into $\chi \chi$, a pair of weakly interacting massive particles (WIMPs) [16, 17], which may explain the existence of dark matter (DM) [18–20]. Such models, in which the DM candidate is a singlet under the gauge symmetries of the SM and the Higgs boson acts as a mediator to the DM, are called Higgs-portal models. They represent a simple extension of the SM to provide a DM candidate and are able to predict the observed relic DM density via $s$-channel $\chi \chi \rightarrow f \bar{f}$ annihilation processes [21], where $f$ denotes a fermion. The phenomenology of such models is very rich, and depends strongly on the spin of the DM candidate.

A long list of searches for invisible Higgs boson decays that target various final states have been performed by the ATLAS and CMS experiments [22–31] at the LHC [32]. The absence of an excess of such events is interpreted as an upper limit on $B_{\text{inv}}$, assuming the SM production cross sections [33] of the Higgs boson. Statistical combinations of search results have been performed, for which ATLAS reported an observed (expected) limit of 0.26 (0.17) [34] and CMS reported 0.19 (0.15) [35] at 95% confidence level (CL), both using Run-1 and up to 36 fb$^{-1}$ of Run-2 data. The results with the highest expected sensitivity are from the vector-boson fusion (VBF) channel, for which ATLAS reported an upper limit of 0.37 (0.28) using 36 fb$^{-1}$ of Run-2 data [27], while CMS reported 0.18 (0.10) using a combination of the full Run-1 and Run-2 data [30].

Global fits to the measurements of visible decay channels of the Higgs boson place indirect constraints on the decay branching fraction $B_{\text{BSM}}$ into final states which are either invisible ($B_{\text{inv}}$) or not included in the combination of visible decays ($B_{\text{undet}}$), such as a BSM Higgs boson that decays into light quarks. A combination of ATLAS and CMS results, using only the cross-section measurements of the visible decay channels from Run-1 data and under the assumption that the couplings between the Higgs boson and the vector bosons are not larger than in the SM, results in an observed (expected) upper limit on $B_{\text{BSM}}$ of 0.34 (0.39) [36]. Direct searches for invisible Higgs boson decays are complementary to indirect constraints from measurements of visible decays. A simultaneous excess in the two approaches would confirm a signal, while a non-zero branching fraction $B_{\text{BSM}}$ in the global fit, but no excess in the direct searches, would point towards other undetected decays. In combinations of direct and indirect constraints, CMS reported a limit on $B_{\text{inv}}$ of 0.22 from 36 fb$^{-1}$ of Run-2 data [37], and ATLAS a limit of 0.30 using up to 80 fb$^{-1}$ of Run-2 data [38].

The data sample used for the analysis reported in this paper corresponds to an integrated luminosity of 139 fb$^{-1}$ of proton-proton ($pp$) collisions at $\sqrt{s} = 13$ TeV recorded by the ATLAS detector. This search targets the VBF production process [39, 40] with its distinct signature of a pair of energetic quark-induced jets with a wide gap in pseudorapidity ($\Delta \eta_{jj}$) resulting in a large invariant mass ($m_{jj}$). Feynman diagrams for the signal and dominant background processes are shown in figure 1. The characteristic VBF topology

---

1 Introduction

This paper presents a direct search for the decay of the Higgs boson [1, 2] into invisible particles. While the expectation for the branching fraction of invisible decays ($B_{\text{inv}}$) from the Standard Model (SM) is 0.12% due to the decay $H \rightarrow ZZ^* \rightarrow \nu \bar{\nu} \nu \bar{\nu}$, several scenarios of physics beyond the Standard Model (BSM) allow for anomalous and sizeable $O(10\%)$ values. In one class of models under consideration [3–15] the Higgs boson might decay into $\chi \chi$, a pair of weakly interacting massive particles (WIMPs) [16, 17], which may explain the existence of dark matter (DM) [18–20]. Such models, in which the DM candidate is a singlet under the gauge symmetries of the SM and the Higgs boson acts as a mediator to the DM, are called Higgs-portal models. They represent a simple extension of the SM to provide a DM candidate and are able to predict the observed relic DM density via $s$-channel $\chi \chi \rightarrow f \bar{f}$ annihilation processes [21], where $f$ denotes a fermion. The phenomenology of such models is very rich, and depends strongly on the spin of the DM candidate.

A long list of searches for invisible Higgs boson decays that target various final states have been performed by the ATLAS and CMS experiments [22–31] at the LHC [32]. The absence of an excess of such events is interpreted as an upper limit on $B_{\text{inv}}$, assuming the SM production cross sections [33] of the Higgs boson. Statistical combinations of search results have been performed, for which ATLAS reported an observed (expected) limit of 0.26 (0.17) [34] and CMS reported 0.19 (0.15) [35] at 95% confidence level (CL), both using Run-1 and up to 36 fb$^{-1}$ of Run-2 data. The results with the highest expected sensitivity are from the vector-boson fusion (VBF) channel, for which ATLAS reported an upper limit of 0.37 (0.28) using 36 fb$^{-1}$ of Run-2 data [27], while CMS reported 0.18 (0.10) using a combination of the full Run-1 and Run-2 data [30].

Global fits to the measurements of visible decay channels of the Higgs boson place indirect constraints on the decay branching fraction $B_{\text{BSM}}$ into final states which are either invisible ($B_{\text{inv}}$) or not included in the combination of visible decays ($B_{\text{undet}}$), such as a BSM Higgs boson that decays into light quarks. A combination of ATLAS and CMS results, using only the cross-section measurements of the visible decay channels from Run-1 data and under the assumption that the couplings between the Higgs boson and the vector bosons are not larger than in the SM, results in an observed (expected) upper limit on $B_{\text{BSM}}$ of 0.34 (0.39) [36]. Direct searches for invisible Higgs boson decays are complementary to indirect constraints from measurements of visible decays. A simultaneous excess in the two approaches would confirm a signal, while a non-zero branching fraction $B_{\text{BSM}}$ in the global fit, but no excess in the direct searches, would point towards other undetected decays. In combinations of direct and indirect constraints, CMS reported a limit on $B_{\text{inv}}$ of 0.22 from 36 fb$^{-1}$ of Run-2 data [37], and ATLAS a limit of 0.30 using up to 80 fb$^{-1}$ of Run-2 data [38].

The data sample used for the analysis reported in this paper corresponds to an integrated luminosity of 139 fb$^{-1}$ of proton-proton ($pp$) collisions at $\sqrt{s} = 13$ TeV recorded by the ATLAS detector. This search targets the VBF production process [39, 40] with its distinct signature of a pair of energetic quark-induced jets with a wide gap in pseudorapidity ($\Delta \eta_{jj}$) resulting in a large invariant mass ($m_{jj}$). Feynman diagrams for the signal and dominant background processes are shown in figure 1. The characteristic VBF topology
Figure 1. Representative Feynman diagrams for (a) the signal and for the three dominant background processes: (b) strong $Z$+jets, (c) electroweak $Z$+jets, and (d) diboson production. The latter two processes are collectively referred to as ‘electroweak’ $Z$ production. Electroweak vertices are shown as red markers. For simplicity, no distinction is made between particles and antiparticles, or between different quark flavours.

provides a powerful way to reject backgrounds from single vector-boson production in association with two jets from QCD radiation, which is the main background after the large backgrounds from multijet processes are suppressed by requiring a large amount of missing transverse momentum ($E_T^{\text{miss}}$). In this analysis, Higgs boson production via the gluon-gluon fusion (ggF) mechanism and in association with a vector boson ($VH$, where $V$ stands for $W$ or $Z$) are also considered as signal, but their contributions are small compared to the VBF process after the selection.

In this analysis, several changes and improvements are made relative to the previous ATLAS paper on this topic [27]. The selection criteria for $m_{jj}$, $\Delta\eta_{jj}$ and the azimuthal angle between the two highest-$p_T$ jets ($\Delta\phi_{jj}$) are relaxed, while the $E_T^{\text{miss}}$ requirement is increased slightly. While the previous analysis rejected events with a third jet, events with three or four jets are accepted in this analysis if these jets are compatible with the hypothesis that they originate from final-state radiation. To profit from the larger data set, a finer binning in $m_{jj}$ and for the first time a binning in $\Delta\phi_{jj}$ ($\Delta\phi_{jj} < 1$ and $1 \leq \Delta\phi_{jj} < 2$) and jet multiplicity are used. This results in improved performance since it better captures the characteristic shape of the signal, which yields higher signal-to-background ratios in kinematic regions with larger values of $m_{jj}$ and at smaller values of $\Delta\phi_{jj}$.

The analysis extracts the signal yield using a binned likelihood fit to 16 search bins and corresponding $Z_{\ell\ell}$ and $W_{\ell\nu}$ ($\ell = e, \mu$) control region bins to estimate the main backgrounds from the $Z(\rightarrow \nu\nu) +$ jets and $W(\rightarrow \ell\nu) +$ jets processes. A recent theoretical calculation [41],
performed in the phase space relevant for this analysis, is used to constrain the ratio of $Z + \text{jets}$ to $W + \text{jets}$ events. This allows a reduction of the statistical uncertainties in the estimate of the background from $Z(\to \nu\nu) + \text{jets}$ events, by also using the $W_\ell\nu$ control region. For the estimation of the multijet background, two independent data-driven techniques are used, and this reduces the overall uncertainties relative to the previously published result. The signal yield extracted from the likelihood fit is used to determine an upper limit on $\mathcal{B}_{\text{inv}}$. The background predictions are validated in a separate kinematic region with two jets and $2 < \Delta\phi_{jj} < 2.5$.

2 ATLAS detector

The ATLAS detector [42] at the LHC covers nearly the entire solid angle around the collision point.\(^1\) It consists of an inner tracking detector (ID) surrounded by a thin superconducting solenoid, electromagnetic (EM) and hadronic calorimeters, and a muon spectrometer incorporating three large superconducting toroidal magnets.

The ID system is immersed in a 2 T axial magnetic field and provides charged-particle tracking in the range $|\eta| < 2.5$. The high-granularity silicon pixel detector covers the vertex region and typically provides four measurements per track, the first hit normally being in the insertable B-layer installed before the start of Run 2 [43–45]. It is followed by the silicon microstrip tracker which usually provides eight measurements per track. These silicon detectors are complemented by the transition radiation tracker, which enables radially extended track reconstruction up to $|\eta| = 2.0$. This detector also provides electron identification information based on the fraction of hits (typically 30 in total) above a higher energy-deposit threshold corresponding to transition radiation.

The calorimeter system covers the pseudorapidity range $|\eta| < 4.9$. Within the region $|\eta| < 3.2$, EM calorimetry is provided by barrel and endcap high-granularity lead/liquid-argon (LAr) calorimeters, with an additional thin LAr presampler covering $|\eta| < 1.8$ to correct for energy loss in material upstream of the calorimeters. Hadronic calorimetry is provided by the steel/scintillator-tile calorimeter, segmented into three barrel structures within $|\eta| < 1.7$, and two copper/LAr hadronic endcap calorimeters. The solid angle coverage is completed with forward copper/LAr and tungsten/LAr calorimeter modules optimised for EM and hadronic measurements respectively.

The muon spectrometer (MS) comprises separate trigger and high-precision tracking chambers measuring the deflection of muons in a magnetic field generated by the superconducting air-core toroids. The field integral of the toroids ranges between 2.0 and 6.0 Tm across most of the detector. A set of precision chambers covers the region $|\eta| < 2.7$ with three layers of monitored drift tubes, complemented by cathode-strip chambers

\(^1\)ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upward. Cylindrical coordinates $(r, \phi)$ are used in the transverse plane, $\phi$ being the azimuthal angle around the z-axis. The transverse momentum of an object is defined as $p_T = \sqrt{p_x^2 + p_y^2}$. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln\tan(\theta/2)$ and the rapidity is given by $y = (1/2) \ln([E + p_z]/[E - p_z])$. Angular separation between two objects is quantified by the distance metric $\Delta R = \sqrt{(\Delta y)^2 + (\Delta\phi)^2}$. 

– 3 –
in the forward region, where the background is highest. The muon trigger system covers the range $|\eta| < 2.4$ with resistive-plate chambers in the barrel, and thin-gap chambers in the endcap regions.

Interesting events are selected to be recorded by the first-level (L1) trigger system implemented in custom hardware, followed by selections made by algorithms implemented in software in the high-level trigger (HLT) [46]. The L1 trigger accepts events from the 40 MHz bunch crossings at a rate below 100 kHz, which the HLT reduces in order to record events to disk at about 1 kHz.

An extensive software suite [47] is used in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

3 Data and simulation

3.1 Data and triggers

The analysis is performed using $pp$ collisions at $\sqrt{s} = 13$ TeV collected by ATLAS in 2015–2018 during stable beam conditions and with all subsystems operational [48]. A total integrated luminosity of 139 fb$^{-1}$ is used, with an uncertainty of 1.7% [49], obtained using the LUCID-2 detector [50] for the primary luminosity measurements.

The data used for this analysis were recorded using several event triggers, as detailed in the following. Signal event candidates were collected with unprescaled $E_\text{miss}$ triggers [51] having the lowest threshold. In 2015, a threshold of 70 GeV was used, which was subsequently raised multiple times to cope with increasing effects from multiple interactions in the same and neighbouring bunch crossings (pile-up) and it reached 110 GeV during the 2017–2018 data-taking period. Control samples for $W$ and $Z$ backgrounds were recorded with the unprescaled single-lepton and dilepton triggers having the lowest $p_T$ thresholds [52, 53].

The thresholds for single-lepton triggers ranged from 20 to 26 GeV for the tightest lepton identification, while for dileptons the thresholds were typically lower, even though the lepton identification was looser. The $E_\text{miss}$ trigger was also used for the muon control regions, and raised the efficiency for muons by 12%–28% depending on the detector region. The $E_\text{miss}$ in the trigger is based only on calorimetric measurements and does not include any reconstructed muons, so the muons behave similarly to invisible particles in this trigger.

One of the techniques to estimate the multijet background uses inclusive jet-triggered data recorded with a set of single-jet triggers with jet $p_T$ thresholds ranging from 15 GeV to $\sim$400 GeV [54]. The single-jet triggers with low $p_T$ thresholds have significant prescale factors that reduce the data collection rate. To compensate for these factors, which can be as large as $O(10^6)$, each event is weighted appropriately. After this treatment, a smooth jet $p_T$ spectrum is obtained.

3.2 Event simulation

Simulated event samples are used to model both the signal and background processes. The full set of simulated samples is summarised in table 1. The generated events were
Table 1. Summary of generators used for simulation. The details and the corresponding references are provided in the body of the text.

<table>
<thead>
<tr>
<th>Process</th>
<th>Generator</th>
<th>ME order</th>
<th>PDF</th>
<th>Parton shower</th>
<th>Tune</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong $V$+ jets</td>
<td>Sherpa 2.2.1 and Sherpa 2.2.7 (mme-filtered)</td>
<td>NLO (up to 2 jets), LO (up to 4 jets)</td>
<td>NNPDF3.0nnlo</td>
<td>Sherpa</td>
<td>Sherpa</td>
</tr>
<tr>
<td>Electroweak $V$+ jets</td>
<td>Herwig 7.2.1</td>
<td>NLO (for 2 jets)</td>
<td>MMHT2014nlo</td>
<td>Herwig angular-order and Pythia 8 dipole recoil</td>
<td>Herwig 7.2</td>
</tr>
<tr>
<td>$V$+ jets $\alpha^2_{\text{ew}}$ interference</td>
<td>MadGraph5_aMC@NLO</td>
<td>LO</td>
<td>PDF4LHC15</td>
<td>Pythia 8</td>
<td></td>
</tr>
<tr>
<td>Strong $VV$+ jets (including $gg \to VV$+jets)</td>
<td>Sherpa 2.2.1 or Sherpa 2.2.2</td>
<td>NLO (up to 1 jet), LO (up to 3 jets)</td>
<td>NNPDF3.0nnlo</td>
<td>Sherpa</td>
<td>Sherpa</td>
</tr>
<tr>
<td>Electroweak $VV$+ jets</td>
<td>Sherpa 2.2.1 or Sherpa 2.2.2</td>
<td>LO</td>
<td>NNPDF3.0nnlo</td>
<td>Sherpa</td>
<td>Sherpa</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>Powheg Box v2</td>
<td>NLO</td>
<td>NNPDF3.0nnlo</td>
<td>Pythia 8</td>
<td>A14</td>
</tr>
<tr>
<td>QCD multijet</td>
<td>Pythia 8.230</td>
<td>LO</td>
<td>NNPDF2.3lo</td>
<td>Pythia 8</td>
<td>A14</td>
</tr>
<tr>
<td>ggF Higgs</td>
<td>Powheg NNLOPS</td>
<td>NNLO</td>
<td>PDF4LHC15</td>
<td>Pythia 8</td>
<td>AZNLO</td>
</tr>
<tr>
<td>VBF Higgs</td>
<td>Powheg</td>
<td>NLO</td>
<td>PDF4LHC15</td>
<td>Pythia 8</td>
<td>AZNLO</td>
</tr>
<tr>
<td>$VH$ Higgs</td>
<td>Powheg Box v2</td>
<td>NLO</td>
<td>PDF4LHC15</td>
<td>Pythia 8</td>
<td>AZNLO</td>
</tr>
</tbody>
</table>

processed through a simulation [55] of the ATLAS detector geometry and response using Geant4 [56], and through the same reconstruction software as the data. The effect of pile-up was modelled by adding detector signals from simulated inelastic $pp$ events to the original hard-scattering event. Those were generated with Pythia 8.186 [57] using the NNPDF2.3lo set of parton distribution functions (PDFs) [58] and the A3 set of tuned parameters (A3 tune) [59]. The lepton energy scales and resolutions, as well as the lepton reconstruction and identification efficiencies, and the trigger efficiencies in the simulation are corrected to match those measured in data.

### 3.2.1 $V$+ jets and $VV$+ jets

The dominant backgrounds in this analysis arise from single weak boson production in association with jets, denoted by $V$+ jets, where the boson decays via $Z \to \nu\nu$ or $W \to \ell\nu$. The $Z \to \ell\ell$ decay modes ($\ell = e$ or $\mu$) are also important in constraining the $Z \to \nu\nu$ contribution. The $V$+jets production processes defined by $\mathcal{O}(\alpha^2_{\text{ew}})$ diagrams, such as in figure 1(b), are referred to as ‘strong’, while ‘electroweak’ $V$ production is defined by $\mathcal{O}(\alpha^4_{\text{ew}})$ diagrams. The latter is divided into diboson and single boson production. Diboson processes (figure 1(d)) can mimic the signal when one of the two vector bosons decays into jets and the other decays into leptons, which are either neutrinos or ‘lost’ during reconstruction. Electroweak single boson production includes the VBF diagram (figure 1(c)), where the jets originate from the initial-state quarks that radiated the vector bosons.
Similar simulation configurations were used for strong production of single-boson \( V + \) jets and for strong and electroweak production of diboson \( VV + \) jets. These are based on the Sherpa 2.2 Monte Carlo (MC) generator [60] with the NNPDF3.0nnlo set of PDFs [61]. Parton-shower matching [62] was performed using either the MEPS@LO [63] or MEPS@NLO [63–66] prescription with associated parameters tuned by the Sherpa authors.

Strong production of \( V + \) jets was simulated using next-to-leading-order (NLO) matrix elements for up to two partons, and leading-order (LO) matrix elements for up to four partons, calculated with the Comix [67] and OpenLoops [68, 69] libraries and the MEPS@NLO prescription. The samples were normalised to a next-to-next-to-leading-order (NNLO) prediction [70]. The simulated \( Z + \) jets events were further weighted as function of \( m_{jj} \), such that the ratio of the strong \( W + \) jets and \( Z + \) jets cross sections matches a dedicated NLO QCD and NLO electroweak calculation in the phase space of a VBF-like signature [41].

The strong \( V + \) jets samples were filtered in dijet mass at the matrix-element level to enhance the statistical precision in the region \( m_{jj} > 0.8 \text{ TeV} \), while generating fewer events in the lower \( m_{jj} \) regions. With the default CSS merging criterion [66] used in Sherpa, the filter was found to be inefficient due to the possibility of selecting different objects before and after the parton shower. To improve the efficiency of the filter, the \( k_t \) algorithm [71, 72] was used as the jet criterion when matching the matrix element to the parton shower. In comparison with the default CSS criterion, the \( k_t \) jet criterion classifies more of the forward partons as being above the threshold to be modelled by the matrix element, giving better agreement between the pre- and post-shower \( m_{jj} \) values.

Electroweak VBF production of \( V + \) jets at NLO was modelled using Herwig 7.2.1 [73, 74] and the VBFNLO [75] external matrix element. The renormalisation and factorisation scales were set to the parton-level \( H_T \) defined as the scalar sum of the \( p_T \) of all parton-level jets with \( p_T > 20 \text{ GeV} \). The MMHT2014nlo [76] PDF set was used for the matrix-element calculation. Two sets of samples were generated, one using the default parton shower with angular ordering, and the other using the dipole shower as an alternative.

Additionally, there is a term of \( \mathcal{O}(\alpha_s^3) \) from interference between the strong and electroweak production modes. This was modelled independently using MadGraph5_aMC@NLO [77]. In the phase space of this search, its contribution was found to be negligible at LO and is therefore not considered further in the background estimation.

Diboson final states \( (VV) \) were simulated at NLO accuracy in QCD for up to one additional parton and at LO accuracy for up to three additional parton emissions. Loop-induced \( gg \rightarrow VV \) processes were generated using LO-accurate matrix elements for up to one additional parton emission. In both cases, the matrix-element calculations were matched and merged with the parton shower using the MEPS@NLO prescription. Electroweak diboson production in association with two jets \( (VVjj) \) was simulated at LO accuracy with the MEPS@LO prescription. The diboson samples include only contributions from doubly resonant diboson diagrams, where either one or both bosons decay leptonically. The off-shell component is assumed to be negligible.
3.2.2 \( t\bar{t}, tW, \) and single top

The production of \( t\bar{t} \), the production of a top quark in association with a W boson (\( tW \)), and single-top \( s \)-channel and \( t \)-channel production were modelled using the Powheg Box [78–81] v2 generator at NLO with the NNPDF3.0nlo PDF set. For \( t\bar{t} \) production, the \( h_{\text{damp}} \) parameter\(^2\) was set to 1.5 \( m_{\text{top}} \) [82]. The events were interfaced to Pythia 8.230 to model the parton shower, hadronisation, and underlying event, with parameters set according to the A14 tune [83] and using the NNPDF2.3lo set of PDFs. The decays of bottom and charm hadrons were performed by EvtGen 1.6.0 [84].

3.2.3 QCD multijet

QCD multijet production was modelled using Pythia 8.230 with leading-order matrix elements for dijet production interfaced to a \( p_T \)-ordered parton shower. The squared renormalisation and factorisation scales were set to the geometric mean of the squared transverse masses of the two outgoing particles, \( \sqrt{(p_{T,1}^2 + m_1^2)(p_{T,2}^2 + m_2^2)} \). The NNPDF2.3lo PDF set was used in the matrix-element generation, the shower, and in the simulation of the multi-parton interactions. The A14 tune was used. Perturbative uncertainties are estimated through event weights [85] that encompass variations of the scales at which the strong coupling constant is evaluated for the initial- and final-state showers as well as the effect of the PDF uncertainty on the shower and the non-singular part of the splitting functions.

3.2.4 Signal

This analysis targets VBF Higgs boson production. The small contributions from the ggF and \( VH \) Higgs boson production processes passing the analysis selection are also treated as signal. The invisible decay of the Higgs boson for ggF, VH, and VBF was simulated using the SM \( H \to ZZ^* \to \nu\bar{\nu}\nu\bar{\nu} \) decay, assuming a branching fraction into this final state of 100%. The difference between this and decays into new invisible particles is negligible, as the Higgs boson is always produced on-shell for any of the scenarios this analysis is sensitive to. Events produced via VBF \( H \to W^*W \) also enter the control regions and were modelled using that same set of generators.

Higgs boson production via gluon-gluon fusion was simulated at NNLO accuracy in QCD using the Powheg NNLOPS program [79–81, 86, 87]. The simulation achieves NNLO accuracy for arbitrary inclusive \( gg \to H \) observables by reweighting the Higgs boson rapidity spectrum from HJ-MtNLO [88–90] to that from HNNLO [91]. The PDF4LHC15nnlo PDF set [92] and the AZNLO tune of Pythia 8 were used. The gluon-gluon fusion prediction from the Monte Carlo samples was normalised to the next-to-NNLO cross section in QCD plus NLO electroweak corrections [33, 93–102].

The \( qq \to VH \) production processes were simulated to NLO accuracy in QCD using the Powheg Box v2 generator with the PDF4LHC15 set of PDFs.

\(^2\)The \( h_{\text{damp}} \) parameter is a resummation damping factor and one of the parameters that controls the matching of Powheg matrix elements to the parton shower and thus effectively regulates the high-\( p_T \) radiation against which the \( t\bar{t} \) system recoils.
Higgs boson production via VBF processes was modelled with Powheg \cite{79–81, 103} and interfaced with Pythia 8 for parton showering, using the dipole recoil scheme, and non-perturbative effects. The Powheg prediction is accurate to NLO and tuned to match calculations with effects due to finite heavy-quark masses and soft-gluon resummations up to next-to-next-to-leading-logarithm accuracy. The PDF4LHC15 PDF set and the AZNLO tune of Pythia 8 were used. The MC prediction was normalised to an approximate-NNLO QCD cross section \cite{104–106}. NLO electroweak (EWK) corrections to VBF Higgs boson ($m_H = 125$ GeV) production were computed separately using HAWK \cite{107}. This was achieved by applying a reweighing factor of $1 - \alpha^{\text{HAWK}}_{\text{NLO EWK}}$ to the Powheg events described above, where $\alpha^{\text{HAWK}}_{\text{NLO EWK}} = -0.000350 \text{ GeV}^{-1} \cdot p_T^{\text{Higgs}} - 0.0430$. This correction was derived in a kinematic region defined by mimicking the analysis selection requirements on jet $p_T$, $\Delta\eta_{jj}$, and $m_{jj}$.

Additional VBF samples with lighter or heavier Higgs bosons were generated in a mass range from 50 GeV to 2 TeV to allow interpretation of the results in scenarios with massive scalar mediators. The parameter values used in the production of these samples are exactly the same as those for the main signal sample; the only difference is the Higgs boson mass. The samples with alternative Higgs boson masses were processed with a fast simulation \cite{55} which relies on a parameterisation of the calorimeter response \cite{108} instead of the full GEANT4-based simulation.

For the background estimation, the normalisation of all Higgs boson samples also accounts for the decay branching fraction calculated with HDECAY \cite{109–111} and Prophecy4F \cite{112–114}.

4 Object and event selection

4.1 Object definitions and event cleaning

Each event is required to have a primary vertex, which is the vertex with the highest sum of squared transverse momenta from at least two associated tracks with $p_T > 0.5$ GeV \cite{115}.

Electrons are reconstructed by matching clustered energy deposits in the EM calorimeter to a track in the ID \cite{116}, including the transition regions between barrel and endcap EM calorimeters at $1.37 < |\eta| < 1.52$. Electron candidates must satisfy $p_T > 4.5$ GeV, $|\eta| < 2.47$, and a ‘loose’ identification criterion. Depending on the $p_T$ and $|\eta|$ range, muons are reconstructed as either an ID track matched to an MS track or track segment(s), an ID track matched to a calorimeter energy deposit compatible with a minimum-ionising particle, or an MS track passing a requirement on compatibility with originating from the IP \cite{117}. Muon candidates must satisfy $p_T > 4$ GeV, $|\eta| < 2.7$, and a ‘very loose’ identification criterion. No isolation requirements are placed on electron and muon candidates.

To be considered as a ‘signal lepton’ stemming from a leptonic decay of a vector boson or a $\tau$-lepton, the leading lepton has to have $p_T > 30$ GeV and fulfil a ‘loose’ isolation criterion. For events with one lepton, i.e. for the control samples used to constrain the background from $W(\rightarrow \ell\nu) + \text{jets}$ processes, the lepton identification uses a ‘tight’ (‘medium’) criterion for electrons (muons) \cite{116, 117}. If two leptons are required, as in the $Z \rightarrow \ell\ell$ control region,
the identification uses ‘loose’ criteria for electrons and muons. Furthermore, the lepton has
to be compatible with originating from the primary vertex.

Photon candidates are reconstructed from clustered energy deposits in the EM
calorimeter [116]. To be considered by the analysis, photons are required to have \( p_T > 20 \) GeV,
to fulfil ‘tight’ identification and isolation criteria and to be within \(| \eta | < 2.37\), excluding the
transition regions between barrel and endcap EM calorimeters. Events that contain an
isolated photon are rejected to ensure orthogonality to other ATLAS analyses [118].

Particle-flow (PFlow) jets are constructed using the anti-\( k_t \) algorithm [119, 120] with a
radius parameter of \( R = 0.4 \), using charged constituents associated with the primary vertex
and neutral PFlow constituents as inputs [121]. Jets are calibrated to the particle scale
using a sequence of corrections, including pile-up subtraction and in situ calibration [122].

Jets are required to have \( p_T > 20 \) GeV and \(| \eta | < 4.5\). For jets with \( p_T < 60 \) GeV and
\(| \eta | < 2.5\), the jet vertex tagger (JVT) discriminant is used to identify jets originating from
the hard-scatter interaction through the use of tracking and vertexing information [123].
The chosen selection criterion accepts hard-scatter jets with an efficiency of about 97%. If
either of the two leading jets has \(| \eta | > 2.5 \) and \( p_T < 120 \) GeV, it must satisfy a threshold
for the forward jet vertex tagger (fJVT) discriminant [124, 125] that has an efficiency for
hard-scatter jets of about 93% and rejects about 58% of pile-up jets with \( p_T > 50 \) GeV. To
further suppress the multijet background for selections with \( E_T^{\text{miss}} < 200 \) GeV a tighter fJVT
criterion has to be satisfied, which has an efficiency of 70%–90% for hard-scatter jets and
rejects more than 80% of pile-up jets depending on the \( p_T \) and \( \eta \) for jets with \( p_T > 50 \) GeV.

Jets containing \( b \)-hadrons (\( b \)-jets) are identified using a multivariate discriminant
(MV2c10) [126]. The chosen criterion is estimated to have 77% efficiency for selecting
\( b \)-jets in an inclusive \( t\bar{t} \) sample, with associated rejection factors of 5 and 134 for jets with charm
and without heavy-flavour content, respectively.

To avoid double counting of energy deposits, the reconstructed objects are required to
be separated in \((y, \phi)\)-space according to the procedure detailed in table 2. For leptons in
the vicinity of jets, the \( \Delta R \) threshold depends on the transverse momentum of the lepton
to account for the collimation of boosted objects.

The missing transverse momentum, with magnitude \( E_T^{\text{miss}} \), is defined as the negative
vectorial sum of the transverse momenta of all selected electrons, muons, photons, and
jets, as well as tracks compatible with the primary vertex but not matched to any of those
objects; this last contribution is called the ‘soft term’ in ATLAS analyses [127, 128]. In the
control samples with one or two signal leptons, the \( E_T^{\text{miss}} \) is adjusted such that these leptons
are treated as invisible particles and referred to as ‘\( E_T^{\text{miss}} \) (without leptons)’. For the \( W_e\nu \)
control regions (see section 5.1), an \( E_T^{\text{miss}} \) significance variable is used to enrich the sample
with multijet events containing a fake electron, i.e. a jet that is misidentified as an electron:

\[
S_{\text{MET}} = \frac{E_T^{\text{miss}}}{\sqrt{p_T^{j1} + p_T^{j2} + p_T^e}},
\]

where the upper indices specify the two highest-\( p_T \) jets (\( j1 \) and \( j2 \)) as well as the electron
\((e)\). To suppress the background from multijet events for which the \( E_T^{\text{miss}} \) stems from
unidentified pile-up jets, or jets mistagged as pile-up, the variable \( p_{T}^{\text{all-jets}} \) is defined as the magnitude of the vectorial sum of the transverse momenta of all jets with \( p_{T} > 20 \text{ GeV} \), including jets tagged as pile-up jets.

Several requirements are applied to suppress non-collision background events [129] with fake jets that do not stem from collisions at the IP. Such fake jets can be caused by electronic noise, and jets from collisions are identified by requiring a good fit to the expected pulse shape for each constituent calorimeter cell. Another source of fake jets are beam-halo interactions with the LHC collimators. Those jets are identified by requirements on their energy distribution in the calorimeters and the fraction of their constituent tracks that originate from the primary vertex. The event is rejected if any selected jet is identified as a fake jet.

### 4.2 Signature from vector-boson fusion

The VBF process exhibits characteristic features which are exploited in this analysis. The two leading jets are typically in opposite hemispheres of the detector and are more forward than jets from non-VBF processes at a comparable parton interaction scale \( \sqrt{s} \). This leads to large values of the pseudorapidity separation \( \Delta \eta_{jj} \) and dijet invariant mass \( m_{jj} \). In QCD multijet events, the two leading jets exhibit a characteristic back-to-back topology \( (\Delta \phi_{jj} \sim \pi) \) in the transverse plane, while for the VBF signal, the Higgs boson tends to have significant \( p_{T} \), which results in smaller values of the azimuthal separation \( \Delta \phi_{jj} \). Hence, a selection based on \( \Delta \phi_{jj} \) is an effective way to suppress the multijet background.

Another distinct feature of the VBF process is the reduced hadronic activity within the rapidity gap between the two leading jets, caused by the absence of colour connection between the two quarks. Therefore, strong \( V+jets \) production can be significantly suppressed by requiring that no further jet with \( p_{T} > 25 \text{ GeV} \) is present. However, initial- and final-state radiation in VBF processes can result in additional jets with \( p_{T} > 25 \text{ GeV} \). Therefore, to enhance the sensitivity of the analysis, events with three or four jets are considered as well, but with additional requirements. The centrality \( C_{i} \) of a third- and fourth-highest-\( p_{T} \) jet

<table>
<thead>
<tr>
<th>Remove</th>
<th>Keep</th>
<th>Matching criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>electron</td>
<td>electron</td>
<td>shared inner-detector track, electron with lower ( p_{T} ) removed</td>
</tr>
<tr>
<td>muon</td>
<td>electron</td>
<td>muon with calorimeter energy deposits and shared inner-detector track</td>
</tr>
<tr>
<td>electron</td>
<td>muon</td>
<td>shared inner-detector track</td>
</tr>
<tr>
<td>photon</td>
<td>electron</td>
<td>( \Delta R &lt; 0.4 )</td>
</tr>
<tr>
<td>photon</td>
<td>muon</td>
<td>( \Delta R &lt; 0.4 )</td>
</tr>
<tr>
<td>jet</td>
<td>electron</td>
<td>( \Delta R &lt; 0.2 )</td>
</tr>
<tr>
<td>electron</td>
<td>jet</td>
<td>( \Delta R &lt; \min(0.4, 0.04 + 10 \text{ GeV}/p_{T}^{e}) )</td>
</tr>
<tr>
<td>jet</td>
<td>muon</td>
<td>number of tracks &lt; 3 and ( \Delta R &lt; 0.2 )</td>
</tr>
<tr>
<td>muon</td>
<td>jet</td>
<td>( \Delta R &lt; \min(0.4, 0.04 + 10 \text{ GeV}/p_{T}^{\mu}) )</td>
</tr>
<tr>
<td>jet</td>
<td>photon</td>
<td>( \Delta R &lt; 0.4 )</td>
</tr>
</tbody>
</table>

**Table 2.** Overview of the overlap removal procedure applied to pairs of objects and the corresponding matching criteria, listed according to priority.
$(i=j3 \text{ and } i=j4)$ [130] is defined as

$$C_i = \exp \left( -\frac{4}{(\eta^1 - \eta^2)^2} \left( \eta^i - \frac{\eta^1 + \eta^2}{2} \right)^2 \right).$$

QCD processes tend to have sub-leading jets emitted close to the midpoint of the rapidity gap, which gives $C_i \approx 1$, while sub-leading jets from final-state radiation in a VBF process tend to have smaller values of $C_i$. Another observable that tests whether a sub-leading jet is compatible with having been radiated from one of the two leading jets in the VBF process is the smaller of the values of the invariant mass of this sub-leading jet and one of the two leading jets, relative to $m_{jj}$:

$$m_i^{\text{rel}} = \frac{\min\{m_{j1,i}, m_{j2,i}\}}{m_{jj}}.$$

Small values of $m_i^{\text{rel}}$ indicate that the additional jet is compatible with final-state radiation.

### 4.3 Event selection

To enter the signal region (SR), the following event requirements must be satisfied:

- The event contains no lepton candidate, nor a photon.
- The event has two, three or four jets with $p_T > 25 \text{ GeV}$ that fulfil the JVT requirement if they have $|\eta| < 2.5$.
- The leading and sub-leading jets have $p_T > 80 \text{ GeV}$ and $50 \text{ GeV}$, respectively. Both jets must fulfil fJVT requirements if they have $|\eta| > 2.5$ and $p_T < 120 \text{ GeV}$.
- To ensure orthogonality to other searches, it is required that not more than one of the jets are $b$-tagged. The fraction of events removed by this requirement is less than 0.02% because most of the jets are outside the acceptance of the tracking detector.
- Any existing third $(i=j3)$ or fourth $(i=j4)$ jet must have $C_i < 0.6$ and $m_i^{\text{rel}} < 0.05$, to be compatible with VBF final-state radiation.
- The event has $E_T^{\text{miss}} > 160 \text{ GeV}$, which strongly suppresses the multijet background.
- The soft term of the $E_T^{\text{miss}}$ is smaller than $20 \text{ GeV}$. This requirement removes events with $W \rightarrow \mu \nu$ decays where the muon is not identified but is reconstructed as a track.
- Further suppression of the multijet background is achieved by requiring $p_T^{\text{all-jets}} > 140 \text{ GeV}$.
- The two leading jets are not back-to-back: $\Delta \phi_{jj} < 2$.
- The two leading jets must fulfil the VBF topology requirements of opposite longitudinal hemispheres ($\eta^1 \cdot \eta^2 < 0$), large pseudorapidity separation ($\Delta \eta_{jj} > 3.8$), and large invariant mass ($m_{jj} > 0.8 \text{ TeV}$).
Figure 2. Schematic view of the 16 bins in the signal region. The shading indicates the signal-to-background ratio, assuming a branching fraction of the Higgs boson to invisible particles of $B_{\text{inv}} = 0.15$. The percentage gives the approximate contribution of signal from invisibly decaying Higgs bosons to each of the bins relative to the total signal yield in the inclusive SR.

For the VBF process, which contributes approximately 90% of the signal events in the SR, the selection efficiency for an invisibly decaying Higgs boson is about 1.1%. For the ggF process, the selection efficiency is two orders of magnitude smaller. The selected events in the SR are separated into 16 bins of different signal purity. Events with exactly two jets and $E_T^{\text{miss}} > 200$ GeV are separated into five $m_{jj}$ bins ($0.8$–$1.0$ TeV, $1.0$–$1.5$ TeV, $1.5$–$2.0$ TeV, $2.0$–$3.5$ TeV, $>3.5$ TeV) and are further separated into two bins with $\Delta \phi_{jj} < 1$ and $1 \leq \Delta \phi_{jj} < 2$. This results in ten SR bins for two-jet events. The highest signal-to-background ratio is obtained for the largest values of $m_{jj}$, while the lower range in $\Delta \phi_{jj}$ also shows better background suppression. Events with three or four jets and $E_T^{\text{miss}} > 200$ GeV form another three bins defined by three $m_{jj}$ ranges: $1.5$–$2.0$ TeV, $2.0$–$3.5$ TeV, and $>3.5$ TeV. Finally, the same three $m_{jj}$ intervals are used to define three more bins with two jets but $160$ GeV $< E_T^{\text{miss}} \leq 200$ GeV. A schematic view of the binning strategy is shown in figure 2.

5 Background estimation

Events from $V+$ jets processes comprise about 95% of the background that enters the SR and their contributions are estimated using dedicated control regions (CRs). A smaller contribution arises from multijet (MJ) processes and is estimated from data. Minor backgrounds from $t\bar{t}$, $VV$, $VVV$, and VBF $H$ production are estimated by MC simulation.

5.1 $V+$ jets background

The backgrounds in the SR from $V+$ jets processes are challenging to model, and hence have large associated theoretical uncertainties. To improve the estimation of these backgrounds,
data-driven technique is implemented that uses CRs defined in kinematic regions analogous to the SR but containing selected $Z(\rightarrow \ell\ell) +$ jets and $W(\rightarrow \ell\nu) +$ jets events. The CRs are divided into 16 bins following the same selection in $m_{jj}$, $\Delta \phi_{jj}$, $E_T^{\text{miss}}$, and $N_{\text{j}\ell}$ as for the SR, but using $E_T^{\text{miss}}$ (without leptons) instead of $E_T^{\text{miss}}$. By construction, the strong and electroweak $V +$ jets processes are probed simultaneously in those CR bins.

The $Z_{\ell\ell}$ CR is based on the same selection criteria as the SR, but the lepton veto is replaced by a requirement of exactly two same-flavour opposite-sign ‘signal leptons’ ($\ell$) with $|m_{\ell\ell} - m_Z| < 25 \text{GeV}$. The $p_T$ threshold for the leading lepton is $30 \text{ GeV}$, while for the sub-leading lepton no further requirement on $p_T$ is applied. To suppress contamination from $VV$ and $VVV$ production, as well as contributions from VBF $H \rightarrow W^*W$ or $H \rightarrow \tau^+\tau^-$, the events are required to have $E_T^{\text{miss}} < 70 \text{ GeV}$, where the leptons are included in the $E_T^{\text{miss}}$ calculation. This exploits the fact that events from those processes are expected to have larger amounts of $E_T^{\text{miss}}$, due to the neutrinos produced by leptonically decaying $W$ bosons or $\tau$-leptons, than events from $Z(\rightarrow \ell\ell) +$ jets processes, for which the $E_T^{\text{miss}}$ arises mainly from jet mismeasurements and peaks at values of $\mathcal{O}(10 \text{ GeV})$. These selection criteria ensure a purity in $Z +$ jets events greater than 90% for all bins. A comparison of the observed yields in the $Z_{\ell\ell}$ CR and the expectation from simulation is shown for selected distributions in figure 3.

The $W$ CRs require exactly one ‘signal lepton’ with $p_T > 30 \text{ GeV}$. The remaining selection criteria are identical to those of the signal region. Events that pass these requirements are separated according to the lepton flavour. The single-electron sample is expected to have a contribution of events with fake electrons originating from multijet processes in which a jet is misidentified as an isolated electron. To suppress this contribution, the events in the single-electron control region ($W_{e\nu}$ CR) are also required to satisfy $S_{\text{MET}} > 4\sqrt{\text{GeV}}$. The single-electron events that fail this $S_{\text{MET}}$ selection define the ‘fake-$e$ CR’, which is enriched with fake electrons since multijet events have no or very little $E_T^{\text{miss}}$ and hence are expected to have low $S_{\text{MET}}$ values, as checked in simulation. In this CR, the fake-electron contribution can be extracted in all 16 bins by subtracting from the data the MC predictions for processes with real electrons. To evaluate the fake-electron background in the $W_{e\nu}$ CR, another control region ($W_{e\nu}$ anti-ID) enriched with fake electrons is defined with the electron passing a loose identification requirement but not the tight one for ‘signal leptons’ (figure 4(a)). For this selection the ratio $R_S$ of the number of data events with high $S_{\text{MET}}$ to the number with low $S_{\text{MET}}$ is calculated after the $W$, $Z$, $VV$, $VVV$ and $t\bar{t}$ backgrounds are subtracted. The average $R_S$ for all 16 bins is $0.12 \pm 0.01$ and its value is evaluated for each $m_{jj}$ range for $160 \text{ GeV} < E_T^{\text{miss}} < 200 \text{ GeV}$ and $E_T^{\text{miss}} > 200 \text{ GeV}$ separately. The fake-electron contribution entering bin $i$ of the $W_{e\nu}$ CR is evaluated as $R_{S,i} \cdot n_{\text{fake-e},i}$, where $n_{\text{fake-e},i}$ is the excess of data events over the expected prompt background in the corresponding bin of the ‘fake-$e$ CR’. This procedure assumes that the identification efficiencies of electrons and $S_{\text{MET}}$ are uncorrelated.

A small contribution of events with fake muons from misidentified jets is expected in the $W_{\mu\nu}$ control region. This contribution is estimated similarly to the fake-electron contribution in the $W_{e\nu}$ CR discussed above. Since events with fake muons tend to contain a significant amount of $E_T^{\text{miss}}$, the variable $S_{\text{MET}}$ is not well suited to increasing the contribution of
Figure 3. Observed and predicted distributions of $m_{jj}$, $\Delta\phi_{jj}$, and $E_T^{\text{miss}}$ (without leptons) in the $Z_{ll}$ control regions, separately for $Z \rightarrow e^+e^-$ ((a), (c), and (e)) and $Z \rightarrow \mu^+\mu^-$ ((b), (d), and (f)). The MC predictions are normalised to cross sections times luminosity and are shown prior to the likelihood fit. Minor contributions from $t\bar{t}$, $VV$, $VVV$, and VBF $H$ with $H \rightarrow \tau^+\tau^-$ or $H \rightarrow W^+W$ are combined and labelled ‘other’. The tiny contributions from $W$+ jets processes are not visible for the given value range. The hatched band indicates statistical and experimental systematic uncertainties. The dielectron and dimuon channels are combined to a dilepton CR in the statistical analysis.
The hashed bands display the combined uncertainty from MC sample sizes and experimental sources.

Figure 4. The figure labelled (a) shows a comparison of the $S_{\text{MET}}$ distribution in data and simulation for the $W_{\mu\nu}$ anti-ID control region, for which the electron passes a loose identification criterion but not the tight one. The figure labelled (b) shows a comparison of the transverse mass $m_T$ distribution in data and simulation for the $W_{\mu\nu}$ anti-ID control region, for which the muon candidate does not pass the loose identification criterion. Minor contributions from $tt$, $VV$, $VVV$, and VBF $H \to \tau^+\tau^-$ or $H \to WW^*$ are combined and labelled ‘other’. The excesses in data, observed at low $S_{\text{MET}}$ and low $m_T$ values, are caused by multijet events containing a fake lepton. The hashed bands display the combined uncertainty from MC sample sizes and experimental sources.
Figure 5. Observed and predicted distributions of $m_{jj}$, $\Delta\phi_{jj}$, and $E_{T}^{\text{miss}}$ (without leptons) in the $W_{\ell\nu}$ [(a), (c), and (e)] and $W_{\mu\nu}$ [(b), (d), and (f)] control regions. The MC predictions are normalised to cross sections times luminosity and are shown prior to the likelihood fit. The small contribution from fake electrons is the expected contribution after the combined likelihood fit. Minor contributions from $t\bar{t}$, $VV$, $VVV$, and VBF $H$ with $H \rightarrow \tau^{+}\tau^{-}$ or $H \rightarrow W^{*}W$ are combined and labelled ‘other’. The hatched band indicates statistical and experimental systematic uncertainties.
Figure 6. The double ratio $R_{TH}^{Z/W} / R_{MC}^{Z/W}$ as a function of generator-level $m_{jj}$, which is used to correct the (a) strong and (b) electroweak $V+\text{jets}$ simulation to achieve state-of-the art precision for the ratio of $Z+\text{jets}$ to $W+\text{jets}$ cross sections in the phase-space of a VBF-like signature as detailed in Reference [41], which is similar to the kinematic region targeted by the analysis. The error bars show the statistical uncertainty of the simulation.

of the analysis. To improve the $Z+\text{jets}$ background estimation, the $W_{\ell\nu}$ CRs are used to determine not only the $W+\text{jets}$ contribution to the SR, but also the $Z+\text{jets}$ background. This is motivated by the similarity of the $W+\text{jets}$ and $Z+\text{jets}$ processes: they have identical QCD couplings for the hadronic part of the process, with only small differences in the mass, spin- and flavour structure. This approach requires an accurate estimate of the ratio $R_{TH}^{Z/W}$ of the $Z+\text{jets}$ and $W+\text{jets}$ cross sections, which is provided by a dedicated NLO QCD and NLO electroweak calculation in the phase space of a VBF-like signature separately for the electroweak and strong processes [41]. The ratio $R_{TH}^{Z/W}$ is evaluated as a function of generator-level $m_{jj}$, and the expectations for $Z+\text{jets}$ processes in all CRs and SRs are reweighted by $R_{TH}^{Z/W} / R_{MC}^{Z/W}$ where $R_{MC}^{Z/W}$ is the $Z$ to $W$ cross-section ratio in the same VBF-like region as is used for the numerator, with ‘MC’ being SHERPA for strong $V+\text{jets}$ samples and HERWIG for electroweak $V+\text{jets}$ samples. The correction to the strong $Z+\text{jets}$ prediction is below 5% across the whole $m_{jj}$ range (figure 6(a)); however, for the electroweak $Z+\text{jets}$ prediction it varies between 20% at low $m_{jj}$ values and 5% at large $m_{jj}$ values (figure 6(b)). The main reasons for these significant corrections at low $m_{jj}$ are missing NLO $s$-channel contributions [131] and missing interference between $s$- and $t$-channel processes in the simulated event samples.

5.2 Multijet background

Events from multijet processes have momentum balance in the transverse plane at generator level, but can have some $E_T^{\text{miss}}$ due to neutrinos produced in subsequent decays of $c$- and $b$-hadrons. At reconstruction level, significant $E_T^{\text{miss}}$ can occasionally be produced by experimental effects, primarily fluctuations in the detector’s response to jets, quantified by the jet energy resolution (JER), but also effects such as very energetic jets not having their energy fully deposited within the calorimeters (‘punch-through’) and misclassification of jets
as originating from the hard-scatter interaction or a pile-up interaction. The signal selection is designed to heavily suppress such events through the requirement of large $E_{\text{T}}^{\text{miss}}$ and small $\Delta \phi_{jj}$. However, in the presence of high pile-up activity a combination of jet mismeasurement and mistagging of pile-up jets can produce signal-like topologies. In such cases, it is possible that the two highest-$p_T$ jets in the event originate from different scattering processes and thus are uncorrelated and can have a small $\Delta \phi_{jj}$ separation. Furthermore, each of the two jets would have its own hadronic recoil system. In most of these cases, the hadronic recoil to the selected jets will result in additional reconstructed jets, and the event will not pass the SR selection. However, some or all of these additional jets might be lost, either because their reconstructed $p_T$ is below the jet $p_T$ requirement or because they are misclassified by the pile-up tagging. The latter effect is a source of $E_{\text{T}}^{\text{miss}}$ that is of particular interest for the investigated signature since one or both of the leading jets are often located outside the tracker acceptance, where the pile-up tagging is less efficient.

The multijet background is estimated using two independent methods: a modified version of the ‘Rebalance and Smear’ (R+S) technique used in previous analyses [27, 132] and a novel data-driven approach, called the ‘pile-up-CR method’, using multijet-enriched control regions. These methods use different approaches and independent data sets to obtain predictions of the multijet background in each SR bin with associated uncertainties. As further detailed below, the final multijet estimate is obtained using either the more precise method, with the other prediction providing validation, or in some parts of the SR a combination of the estimates from the two methods. An overview of all control regions used for the estimation and validation of the multijet background is presented in table 3.

The R+S method has two distinct steps. First, a multijet event sample is constructed in which each event exhibits momentum balance in the transverse plane and hence approximates the particle level. Next, the detector response and event reconstruction is emulated multiple times per event, which will occasionally produce events that fulfil the SR requirements. The R+S method is performed twice using different input events, namely single-jet-triggered data events and simulated multijet events. These approaches make different assumptions and result in two separate predictions. Each jet is classified as either a hard-scatter (HS) jet or a pile-up (PU) jet. When simulated events are used, a jet is identified as a HS jet if it is matched with $\Delta R < 0.1$ to a generator-level jet from the hard-scatter process, while jets that are not matched are classified as PU jets. Events are only considered further if the transverse momentum imbalance calculated from all matched generator-level jets is less that 35 GeV. This ensures that all significant HS jets have been identified. The subset of events with two or more HS jets but no PU jet with $p_T > 50$ GeV defines a HS-only sample, while the HS+PU sample contains events with two or more HS jets and at least one PU jet with $p_T > 50$ GeV. When data events are used, HS jets cannot be identified by a $\Delta R$ matching to generator-level jets. Instead, information from the pile-up tagging is used. Those with no pile-up-tagged jets with $p_T > 25$ GeV define the HS-only sample. PU jets stemming from the same scattering process are identified by requiring two pile-up-tagged jets with $p_T > 25$ GeV, which also fulfil the requirements that their average $p_T$ is above 40 GeV and that they are back-to-back ($\Delta \phi_{jj} > 2.7$). Events with exactly two such jets define the HS+PU sample. If simulated events are used as input, the advantage is a clean
\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
Region & Definition & Binned & Purpose \\
\hline
MJ CR & 100 GeV < $E_{\text{miss}}^T$ < 200 GeV and $m_{jj}$ > 400 GeV & no & Determine HS-only and HS+PU composition \\
 & $N_{\text{jet}} = 2$, $\Delta_{\phi_{jj}} > 2.5$ & & \\
 & no requirement on $p_{\text{T}}^{\text{all-jets}}$ & & \\
\hline
low-$m_{jj}$ CR & 150 GeV < $E_{\text{miss}}^T$ < 200 GeV & no & Multijet (R+S) normalisation \\
 & 2 \leq N_{\text{jet}} \leq 4, 200 GeV < $m_{jj}$ < 800 GeV & & Estimate non-closure uncertainty \\
 & $\Delta_{\phi_{jj}} > 2.5$, ($|y_{j1}| > 2.4$ or $|y_{j2}| > 2.4$) & & \\
 & $\Delta_{\delta_{jj}} < 1$, no requirement on $p_{\text{T}}^{\text{all-jets}}$ & & \\
\hline
mid-$m_{jj}$ CR & 160 GeV < $E_{\text{miss}}^T$ < 200 GeV & no & Same as low-$m_{jj}$ CR \\
 & $N_{\text{jet}} = 2$, 800 GeV < $m_{jj}$ < 1500 GeV & & \\
 & $\Delta_{\phi_{jj}} < 1$, no requirement on $p_{\text{T}}^{\text{all-jets}}$ & & \\
\hline
pile-up CR & |$y_{|j}|$ > 2.4 and j1 is PU tagged & yes & Scaled by $R_{\text{MJ}}$ to obtain multijet yield in SR \\
 & no pile-up-tagging requirement for j2 & & \\
 & 2 \leq N_{\text{jet}} \leq 4 & & \\
 & no $m_{\text{rel}}$ requirement for 1 < $\Delta_{\phi_{jj}} < 2$ & & \\
\hline
low-MET CR & 100 GeV < $E_{\text{miss}}^T$ < 160 GeV & yes & Numerator of $R_{\text{MJ}}$ \\
\hline
low-MET PU-CR & same as pile-up CR, but 100 GeV < $E_{\text{miss}}^T$ < 160 GeV & yes & Denominator of $R_{\text{MJ}}$ \\
\hline
low-MET VR & 110 GeV < $E_{\text{miss}}^T$ < 150 GeV & no & Performance check of pile-up CR method \\
 & loose JVT requirement & & \\
 & no requirement on $p_{\text{T}}^{\text{all-jets}}$ & & \\
\hline
\end{tabular}
\caption{Control regions used for the multijet background estimation. Only changes relative to the SR selection are given. The third column shows if a the control region is binned in $\Delta_{\phi_{jj}}$ and $m_{jj}$ like the signal region. Except for the MJ CR, which uses single-jet-triggered data, all control regions use $E_{\text{miss}}^T$-triggered data.}
\end{table}

The HS jets are ‘rebalanced’, i.e. the momenta of the jets are adjusted within their experimental resolution such that no transverse momentum imbalance is present. If data events are used as input, the method relies on pile-up tagging to identify the PU jets, which is less efficient. But in the latter case, the method takes the pile-up contribution from data, which is expected to be more accurate than if it were taken from simulation.

The HS jets are ‘rebalanced’, i.e. the momenta of the jets are adjusted within their experimental resolution such that no transverse momentum imbalance is present. If data events are used as input, the method relies on pile-up tagging to identify the PU jets, which is less efficient. But in the latter case, the method takes the pile-up contribution from data, which is expected to be more accurate than if it were taken from simulation.

The HS jets are ‘rebalanced’, i.e. the momenta of the jets are adjusted within their experimental resolution such that no transverse momentum imbalance is present. If data events are used as input, the method relies on pile-up tagging to identify the PU jets, which is less efficient. But in the latter case, the method takes the pile-up contribution from data, which is expected to be more accurate than if it were taken from simulation.

The HS jets are ‘rebalanced’, i.e. the momenta of the jets are adjusted within their experimental resolution such that no transverse momentum imbalance is present. If data events are used as input, the method relies on pile-up tagging to identify the PU jets, which is less efficient. But in the latter case, the method takes the pile-up contribution from data, which is expected to be more accurate than if it were taken from simulation.

The HS jets are ‘rebalanced’, i.e. the momenta of the jets are adjusted within their experimental resolution such that no transverse momentum imbalance is present. If data events are used as input, the method relies on pile-up tagging to identify the PU jets, which is less efficient. But in the latter case, the method takes the pile-up contribution from data, which is expected to be more accurate than if it were taken from simulation.
of each event are rotated relative to each other by a random azimuthal angle multiple times. Following this procedure, the event selection is applied to populate the SR and CR bins.

Independently of how the HS-only and the HS+PU samples are defined, the two multijet background components are normalised by a fit to the $\Delta \phi_{jj}$ distribution in a loose multijet control region using single-jet-triggered data (MJ CR, defined in table 3). The $\Delta \phi_{jj}$ distribution is well suited for this, since HS-only topologies are expected to accumulate at large $\Delta \phi_{jj}$, while at low $\Delta \phi_{jj}$ the HS+PU topology should be dominant. The shapes of the $m_{jj}$ and $\Delta \phi_{jj}$ distributions after the normalisation are shown in figure 7 for the case of the HS-only and HS+PU samples being derived from simulated events.

The final step in the multijet background estimation is to apply a correction to account for the inefficiency of the $E_T^{\text{miss}}$ triggers used to collect the data in the SR bins. For $E_T^{\text{miss}}$ values above 160 GeV, the $E_T^{\text{miss}}$ triggers are almost fully efficient and the corrections are small compared with the other uncertainties, which are summarised later. The multijet background is estimated separately for three data-taking periods (2015–2016, 2017, and 2018) to account for different pile-up, beam, and trigger conditions.

The multijet prediction from the R+S method is tested in two dedicated control regions. The first (low-$m_{jj}$ CR) is defined at low $m_{jj}$ (200 GeV < $m_{jj}$ < 800 GeV) for $E_T^{\text{miss}}$ values between 150 and 200 GeV. The selection criterion for $\Delta \eta_{jj}$ is relaxed to $\Delta \eta_{jj} > 2.5$ but one of the two leading jets is required to be forward with $|\eta| > 2.4$. Another control region (mid-$m_{jj}$ CR) is defined at intermediate $m_{jj}$ (800 GeV < $m_{jj}$ < 1500 GeV) for $E_T^{\text{miss}}$ values between 160 and 200 GeV and $N_{\text{jet}} = 2$. Both control regions have no requirement on $p_T^{\text{all-jets}}$ and furthermore the events have to fulfil $\Delta \phi_{jj} < 1$ to increase each sample’s purity in multijet
events. The remaining selection requirements are the same as for the SR, i.e. the used data
are collected by $E_T^{\text{miss}}$ triggers in contrast to the MJ CR, which uses single-jet triggers.
For each of the two control regions, the multijet prediction is normalised to the observed
data. The average normalisation factor from the two multijet control regions is used for
the nominal multijet prediction, while half of the difference between the two normalisation
factors is assigned as a ‘non-closure’ uncertainty of the method.

Further systematic uncertainties in the R+S multijet prediction arise from three sources.
The smearing is repeated with systematically varied jet energy response distributions to
account for potential differences between data and simulation. Two uncertainty variations
of the jet energy response are considered: a 15% widening of its ‘core’ part, which is
modelled by a Gaussian distribution, and a 50% increase of the non-Gaussian tails. These
uncertainties are motivated by finely binned data-MC comparisons of the widths and the
non-Gaussian tails of the dijet $p_T$ asymmetry, which is closely related to the calorimeter’s
response to jets [122]. The differences from the nominal prediction are used as an additional
systematic uncertainty. The prediction of the multijet background is made with seed
events from simulation and from data, and the resulting difference is used as a further
systematic uncertainty. Table 4 summarises the inclusive multijet predictions and the
assigned systematic uncertainties.

An independent and novel approach used to predict the multijet background is the
pile-up-CR method. A control region enriched with multijet events (PU-CR) is defined by
inverting the fJVT criterion for the leading jet, i.e. by requiring that the leading jet in $p_T$
is identified as a forward pile-up jet. No requirement is made on the fJVT value of the
sub-leading jet in $p_T$. This ensures that the PU-CR contains events with both leading jets
originating from different interactions, as well as events in which they do originate from
the same scattering process. This control region’s purity in multijet events is around 95%
for $E_T^{\text{miss}}$ larger than 200 GeV and greater than 92% for $E_T^{\text{miss}}$ between 160 and 200 GeV.
The pile-up CR is binned in $\Delta \phi_{jj}$, $m_{ij}$, and $E_T^{\text{miss}}$ bins corresponding to the signal region
bins. The event yield $N^i_{\text{PU-CR}}$ in the pile-up CR bin $i$ is scaled to the expected multijet
background in the corresponding SR bin by a normalisation factor $R_{\text{MJ},i}$, which is measured
at lower $E_T^{\text{miss}}$ where the multijet background is a significant contribution. Because forward
pile-up jets are identified only up to $p_T$ values of 120 GeV, the yields in the pile-up CRs are
biased to lower values in $E_T^{\text{miss}}$ compared with the SR if only events with two jets are used.
To ensure that the normalisation factor is valid over the considered range in $E_T^{\text{miss}}$ the event
selection for the pile-up CR allows up to four jets, following the jet requirements on $C_i$
and $m_i^{\text{rel}}$ for $i = 3$ or 4 as defined in section 4.3. Because not many events contribute to the
pile-up CR bins for $1 < \Delta \phi_{jj} < 2$ the requirement on $m_i^{\text{rel}}$ for a possible third or fourth jet is
dropped. The normalisation factor $R_{\text{MJ},i}$ is calculated from the event yields $N^i_{\text{low-MET}}$ and
$N^i_{\text{PU-CR,low-MET}}$ at low $E_T^{\text{miss}}$ between 100 GeV and 160 GeV with the remaining selection
criteria being identical to the SR and pile-up CR bin definitions, respectively:

$$R_{\text{MJ},i} = \frac{N^i_{\text{low-MET}} - B^\text{low-MET}_{\text{non-MJ},i}}{N^i_{\text{PU-CR,low-MET}} - B^\text{PU-CR,low-MET}_{\text{non-MJ},i}}.$$ 

Here, $B^\text{low-MET}_{\text{non-MJ},i}$ and $B^\text{PU-CR,low-MET}_{\text{non-MJ},i}$ are the contributions from non-multijet processes to
the corresponding bins. The normalisation factor is then applied to the event yield in the pile-up CR after subtraction of the non-multijet backgrounds to obtain a prediction of the multijet yield in the SR:

\[ B_{\text{MJ},i}^{\text{SR}} = R_{\text{MJ},i} \cdot \left(N_{\text{PU-CR}}^{\text{MJ}} - B_{\text{PU-CR,non-MJ}}^{\text{CR}}ight), \]

with the non-multijet backgrounds \( B_{\text{PU-CR,non-MJ}}^{\text{CR}} \) taken from simulation but normalised to the data in appropriate control regions that have no requirement on the fJVT value of the leading jet. For SR bins 14 to 16 (160 GeV < \( E_{\text{T}}^{\text{miss}} \) < 200 GeV), the ratio \( R_{\text{MJ},i} \) and the pile-up CRs use the tighter fJVT criterion than for the remaining SR bins as mentioned in section 4.1. The stability of \( R_{\text{MJ},i} \) as a function of \( E_{\text{T}}^{\text{miss}} \) was checked in bins of \( E_{\text{T}}^{\text{miss}} \) with a width of 10 GeV between values of 100 GeV and 160 GeV for the SR bins with 160 GeV < \( E_{\text{T}}^{\text{miss}} \) < 200 GeV and between values of 100 GeV and 200 GeV for SR bins with \( E_{\text{T}}^{\text{miss}} \) > 200 GeV. To increase the statistical precision of this test, the bins 1–5, 6–10 and 14–16 are merged. No statistically significant dependence on \( E_{\text{T}}^{\text{miss}} \) is observed and an uncertainty of 20% is assigned for this effect, which corresponds to the statistical precision of the test. For the individual bins, the statistical uncertainty of the measured \( R_{\text{MJ},i} \) is added in quadrature to this value, resulting in an uncertainty of 20% to 27% for the normalisation factor. The method is validated by using a selection enriched in multijet events (low-MET VR) with a reduced \( E_{\text{T}}^{\text{miss}} \) requirement (110 GeV < \( E_{\text{T}}^{\text{miss}} \) < 150 GeV), no requirement on \( p_{\text{T}}^{\text{all-jets}} \) and the same loose fJVT requirement as used for the SR bins with \( E_{\text{T}}^{\text{miss}} \) > 200 GeV (figure 8). The contributions from \( V+ \) jets processes are normalised in appropriate CRs and the data agrees with the background prediction within the uncertainties.

**Figure 8.** The distribution of (a) \( \Delta\phi_{jj} \) and (b) \( m_{jj} \) in the low-\( E_{\text{T}}^{\text{miss}} \) validation region. The multijet background is predicted by the pile-up-CR method. Minor contributions from \( t\bar{t},VV,\) and \( \text{VBF } H \) with \( H \rightarrow \tau^+\tau^- \) or \( H \rightarrow W^*W^* \) are combined and labelled ‘other’. The hatched band indicates the statistical and systematic uncertainties of the backgrounds.
Table 4. Predicted multijet background, before the likelihood fit, entering the signal region bins with $N_{\text{jet}} = 2$ with associated uncertainties from the Rebalance and Smear method and the pile-up-CR method. The quoted uncertainty in the inclusive prediction is estimated by summing uncertainties from all sources in quadrature over the SR bins. The individual sources of uncertainty are given for the inclusive SR but also as a range for the individual SR bins. The range for the statistical uncertainties of the pile-up-CR method are only shown for bins 1 to 5; for bin 6 to 10 no bin-by-bin uncertainties are evaluated since those are merged, and for bins 14 to 16 the statistical uncertainties are negligible.

The predicted yields and their uncertainties are summarised in table 4 for SR bins with $N_{\text{jet}} = 2$. The SR bins with $3 \leq N_{\text{jet}} \leq 4$ are excluded from the comparison, because for those no predictions from the pile-up-CR method are available. The multijet background predictions from the two methods for the inclusive signal region agree within their uncertainties. Since for the SR bins with $\Delta \phi_{jj} < 1$ and the bins with $160 \text{ GeV} < E_{\text{T}}^{\text{miss}} < 200 \text{ GeV}$ the uncertainties in the expected yields from the pile-up-CR method are smaller than those from the R+S method, the former are used. Because not many events contribute to the pile-up CR bins for $1 < \Delta \phi_{jj} < 2$ associated with the SR bins 6–10, the prediction with the pile-up-CR method is made inclusively for these bins and the result is then distributed over the individual bins in accord with predictions from the R+S method. For the SR bins with $3 \leq N_{\text{jet}} \leq 4$ the multijet prediction from the R+S approach is adopted in order to avoid using the same pile-up-CR bins as used for SR bins with $N_{\text{jet}} = 2$.

6 Systematic uncertainties

6.1 Theoretical uncertainties

The theoretical uncertainties affect both the signal and background processes and originate from the limited order at which the matrix elements are calculated, the matching of those calculations to parton showers, and the uncertainty of the proton PDFs. For the minor backgrounds from $tt$, $VV$, $VVV$, and VBF $H$ production with $H \rightarrow \tau^+ \tau^-$ or $H \rightarrow W^*W$ the theoretical uncertainties have been found to be negligible.
6.1.1 $W$ + jets and $Z$ + jets uncertainties

The higher-order matrix-element effects and parton shower matching uncertainties are assessed by varying the four scale choices used in the event generation: the renormalisation, factorisation, resummation, and CKKW matching scales [133].

The factorisation and renormalisation scales are varied upwards and downwards by a factor of two using on-the-fly varied event weights in the SHERPA MC simulation. The use of event weights reduces the statistical uncertainty of the estimate compared with using independently generated samples. The corresponding uncertainties are calculated by taking an envelope of the seven factorisation/renormalisation scale variations: the central values, each scale varied separately, and both scales varied coherently. For the strong $V$ + jets background, the effect of the seven-point scale variations on the expected yield in bins of $m_{jj}$ are $+27\%$ $-18\%$ at low $m_{jj}$, increasing to $+43\%$ $-26\%$ at high $m_{jj}$. The corresponding values for the electroweak $V$ + jets background are $2\%$–$14\%$ for $Z$ + jets and $1\%$–$8\%$ for $W$ + jets.

The resummation scale is also varied by a factor of two, while the CKKW scale is decreased to $15 \text{ GeV}$ and increased to $30 \text{ GeV}$ from the nominal $20 \text{ GeV}$ value. These are applied to the samples using event-by-event reweighting factors calculated with SHERPA 2.1, thereby reducing the overall CPU resources needed to generate the required samples. The reweighting factors are binned in $p_T^V$, the transverse momentum of the produced vector boson, and the number of generator-level jets in the event. The relative error, calculated as the average of the relative effects of the upwards and downwards variations, ranges from $4\%$ to $8\%$ for the resummation scale and from $4\%$ to $6\%$ for the CKKW scale.

The effects of the PDF uncertainties on the $W$ and $Z$ backgrounds are evaluated separately for each bin as the standard deviation of the yields obtained by using the different PDF replicas of the NNPDF set. This results in uncertainties ranging from $1\%$ to $2\%$.

All theoretical uncertainties are treated as uncorrelated between strong and electroweak processes; however, uncertainty sources affecting $Z$ + jets and $W$ + jets events from the same type of process are treated as correlated. This means that an uncertainty variation that increases the strong $Z$ + jets background in the SR also increases the strong $W$ + jets background in the SR and the corresponding backgrounds in the $Z$ and $W$ CRs. This leads to a significant decrease in the uncertainty of the total background yield in each signal region bin. To quantify how well the effects of these uncertainties cancel out, the variation of the ratio of each $V$ + jets background in the SR to that in the corresponding CR is calculated. The uncertainty in the value of each ratio is $1\%$–$3\%$ and is dominated by the statistical uncertainties of the samples used.

The $V$ + jets prediction is estimated by the likelihood fit (see section 7) with one normalisation factor per SR bin, so $Z$ + jets and $W$ + jets backgrounds are assumed to be correlated. Several uncertainties are then assessed by a dedicated NLO QCD and NLO electroweak calculation [41] of the $R_{Z/W}$ ratio, which is used to refine the MC prediction of the $W$ + jets and $Z$ + jets backgrounds (see section 5.1). The relative size of the impact of the NLO QCD corrections on the $R_{Z/W}$ ratio is taken as the uncertainty, which amounts to $\sim 2\%$ for both the strong and electroweak processes across $m_{jj}$ bins. The QCD and electroweak corrections are applied either multiplicatively (default) or additively, resulting
Figure 9. Fractional uncertainties of the double ratio $R_{Z/W}^{TH}/R_{Z/W}^{MC}$. ‘QCD’ refers to the change in the ratio induced by the NLO QCD corrections, ‘Mix’ shows the difference between additive and multiplicative application of the NLO QCD and NLO electroweak corrections, and ‘PS’ stands for the uncertainties attributed to the parton shower. For the strong $V+$ jets processes (a), ‘Sherpa jet veto’ denotes the difference between the ratios with and without a jet veto being applied. For the electroweak $V+$ jets processes (b), the three ‘$\Delta \phi_{jj}$’ uncertainty components quantify the impact of the $\Delta \phi_{jj}$ dependence on the correction and are evaluated in the ranges 0–1, 1–2, and >2.

in a difference of less than 1% in all bins, which is assigned as an uncertainty. For the strong process, parton shower uncertainties of 3% to 5% are calculated as the difference between $R_{Z/W}$ values at NLO with and without parton shower simulation. For the electroweak process, they are calculated as the difference between dipole and angular ordering in HERWIG, resulting in uncertainties of less than 2%. The dedicated NLO QCD and NLO electroweak calculation does not use the same VBF selection as the analysis, nor does it give an $R_{Z/W}$ value for each CR and SR bin. Instead, for strong $V+$ jets production the $R_{Z/W}$ differences between the VBF-like region used to define $R_{Z/W}$ and the individual CR and SR regions are modelled with SHERPA. The dominant effect is due to the jet-veto requirement, which is estimated from SHERPA as the change in $R_{Z/W}$ when applying the jet veto. It ranges from 2% at low $m_{jj}$ to 13% at high $m_{jj}$, where the contribution from strong $V+$ jets production is comparatively small. For electroweak $V+$ jets production, the dominant effect is the difference between the $R_{Z/W}$ values for the inclusive $\Delta \phi_{jj}$ range and the individual generator-level $\Delta \phi_{jj}$ ranges of 0–1, 1–2, and >2. This is estimated from the dedicated NLO QCD and NLO electroweak calculation and ranges up to 11%. The difference between the requirement on $\Delta \eta_{jj}$ for the SR definition ($\Delta \eta_{jj} > 2.5$) and that for the NLO calculation ($\Delta \eta_{jj} > 2.5$) is covered by the assigned uncertainties and has a negligible impact on the final results. A summary of the theoretical uncertainties of $R_{Z/W}$ is given in figures 9(a) and 9(b) for the strong and electroweak $V+$ jets processes, respectively.

6.1.2 Signal uncertainties

The inclusive VBF and ggF Higgs boson production cross sections and uncertainties are provided by the LHC Higgs working group [33]. For the VBF process, there is an uncertainty
due to the $p_T$-dependent NLO electroweak correction. This is estimated from HAWK to be 2%. The $m_{jj}$-dependent effects of the renormalisation and factorisation scale uncertainties and their correlations were computed by the LHC Higgs working group [134] for the POWHEG+PYTHIA signal sample. The factorisation and renormalisation scales are varied independently. Additionally, the Stewart-Tackmann technique [135] is used to compute the uncertainty due to the jet veto. These uncertainties are in the range from 1% to 3%. By comparison with a POWHEG+HERWIG 7 sample, parton shower uncertainties are estimated to have a 2%–4% effect on the VBF signal. PDF uncertainties assessed from the NNPDF set affect the VBF signal by 1%–2%.

Uncertainties associated with the ggF process are also evaluated by renormalisation and factorisation scale variations. A Stewart-Tackmann procedure is used to estimate the jet bin migration uncertainty. This gives a 45% uncertainty in bins with $N_{\text{jet}} = 2$ and 41% in the bins with $N_{\text{jet}} = 3$ or 4. The smaller PDF and parton shower uncertainties are also included.

### 6.2 Experimental uncertainties

Several experimental uncertainties impact the sensitivity of the analysis. They can be grouped into three categories: uncertainties in the luminosity, uncertainties in the trigger efficiencies, and uncertainties related to the used physics objects such as electrons, muons, jets, and the $E_T^{\text{miss}}$.

The uncertainty in the luminosity is 1.7% [49] and impacts only the signal yield because its effect on the dominant $V+$jets backgrounds cancels out in the chosen approach (see section 7).

To maximise the sensitivity of the analysis, the $E_T^{\text{miss}}$ requirement of at least either 160 or 200 GeV for the signal selection is chosen. But since at these values the $E_T^{\text{miss}}$ triggers are yet not fully efficient, systematic uncertainties have to be considered to account for possible trigger efficiency differences between data and simulation. This is done by comparing the combined L1+HLT efficiency turn-on curve for simulated signal events with the turn-on curve derived from data using a sample recorded with single-muon triggers. They agree within the statistical uncertainty, which is below 1%. To cover the difference between the trigger efficiencies for $W \rightarrow \mu \nu$ and $Z \rightarrow \nu \nu$, a further systematic uncertainty is taken from the difference found in simulated samples, which is smaller than 0.5% for all signal region bins. For the $W$ and $Z$ control regions, scale factors and uncertainties of the corresponding single-lepton and dilepton triggers are applied [52, 53].

For electrons and muons, uncertainties in the reconstruction and isolation efficiencies [136, 137], as well as in energy scale and resolution [116, 138], are considered. For the electron veto, an uncertainty in the electron inefficiency is taken into account. For jets, uncertainties in the energy scale and resolution [139] and in the pile-up tagging efficiencies are used [123, 125]. The above uncertainties associated with the reconstructed objects are propagated to the calculation of $E_T^{\text{miss}}$ to evaluate their impact on the predicted yields. The uncertainties in the scale and resolution of the $E_T^{\text{miss}}$ soft term are also propagated to the overall $E_T^{\text{miss}}$ uncertainty [128].
7 Statistical model

A simultaneous maximum-likelihood fit is used to extract the signal yield from the observed event counts in the signal region and control regions. The likelihood is a product of terms for the Poisson probabilities to observe each data count of bin \( i \), \( \mathcal{N}^{\text{CR/SR}}_i \), from a signal-plus-background model. Furthermore, it depends on a set of parameters representing the yields from multijet processes in the fake-\( e \), fake-\( \mu \) and pile-up CRs, and on a set of nuisance parameters \( \theta_j \) that represent experimental and theoretical uncertainties. The strong and electroweak components of the \( V + \) jets estimates in each SR bin \( i \) and the corresponding CR bins are scaled by one dedicated free parameter \( \beta_i \). Since the \( V + \) jets CR is enriched in events from \( V + \) jets processes, the \( \beta_i \) factors are almost exclusively determined so as to match the data counts in these CR bins, and these normalisation constraints are appropriately transferred to the \( V + \) jets prediction in the corresponding SR bin.

The dependence of all background \((B)\) and signal \((S)\) yields on the nuisance parameters \( \vec{\theta} \) is omitted to simplify the notation. The signal region yield likelihood is

\[
\mathcal{L}^{\text{SR}} = \prod_i \mathcal{P} \left( \mathcal{N}^{\text{SR}}_i \mid \beta_i \cdot B^{\text{SR}}_{Z,i} + \beta_i \cdot B^{\text{SR}}_{W,i} + B^{\text{SR}}_{\text{MJ},i} + B^{\text{SR}}_{\text{other},i} + \mu \cdot S^{\text{SR}}_i \right),
\]

where \( B^{\text{SR}}_{W,i} \) and \( B^{\text{SR}}_{Z,i} \) are the \( W + \) jets and \( Z + \) jets yields in the simulation, with the latter already weighted by the double ratio \( \mathcal{R}^{Z/W}_{\text{TH}} / \mathcal{R}^{Z/W}_{\text{MC}} \) as discussed in section 5.1. The \( B^{\text{SR}}_{\text{MJ},i} \) are the multijet background estimates detailed below; \( B^{\text{SR}}_{\text{other},i} \) are sums of the simulated \( t\bar{t} \), \( VV \), \( VVV \) yields and VBF \( H \) with \( H \to \tau^+\tau^- \) or \( H \to WW \); \( S^{\text{SR}}_i \) is the simulated yield for an invisible Higgs boson decay branching fraction of 100%; and \( \mu \) is the signal strength. Since the same \( \beta_i \) is used for \( W + \) jets and \( Z + \) jets contributions, the larger number of events in the \( W_\ell\nu \) CRs reduces the uncertainty of the \( Z + \) jets background, particularly in bins with high \( m_{jj} \) requirements.

The background from multijet events, \( B^{\text{SR}}_{\text{MJ},i} \), is determined as detailed in section 5.2 by two independent methods. The pile-up-CR method uses a multijet-enriched control region that is binned according to the SR to constrain the free fit parameters \( n_{\text{PU-MJ},i} \). Those are scaled by the ratio \( \mathcal{R}_{\text{MJ},i} \) to give the final multijet estimates. In bins 6–10, which have high \( E_{\text{miss}} \) and \( \Delta \phi_{jj} < 2 \), the total event yield is evaluated using the merged bins of the PU-CR, while the shape, i.e. the fraction \( f_{R+S,i} \) of events in each bin, is predicted by the R+S method. For bins 11–13, with \( N_{\text{jet}} > 2 \), the prediction is obtained solely by using the R+S method. The corresponding likelihood is

\[
\mathcal{L}^{\text{PU-CR}} = \prod_{i \in [1,5] \text{ or } [6–10] \text{ or } [14,16]} \mathcal{P} \left( \mathcal{N}^{\text{PU-CR}}_i \mid B^{\text{PU-CR}}_{\non\text{MJ},i} + n_{\text{PU-MJ},i} \right),
\]

where \( B^{\text{PU-CR}}_{\non\text{MJ},i} \) is the non-multijet contribution determined from simulation, and \{6–10\} represents the index of the merged bins. The full multijet estimate that enters the signal region likelihood is

\[
B^{\text{SR}}_{\text{MJ},i} = \begin{cases} n_{\text{PU-MJ},i} \cdot R_{\text{MJ},i} & \text{for } i \in [1,5] \text{ or } [14,16] \\ n_{\text{PU-MJ},i\{6–10\}} \cdot R_{\text{MJ},i\{6–10\}} \cdot f_{R+S,i} & \text{for } i \in [6,10] \\ B^{\text{SR}}_{\text{MJ-R+S},i} & \text{for } i \in [11,13] \end{cases}
\]
The \( V+\text{jets} \) backgrounds are constrained by dedicated \( W \) and \( Z \) control regions and the associated likelihood is:

\[
\mathcal{L}^{V+\text{jets-CR}} = \prod_i P \left( N_i^{Z\text{CR}} \mid \beta_i \cdot B_{Z,i}^{Z\text{CR}} + B_{\text{non}-Z,i}^{Z\text{CR}} \right) \\
\times \prod_i P \left( N_i^{W\mu\text{CR}} \mid \beta_i \cdot B_{W,i}^{W\mu\text{CR}} + B_{\text{non}-W,i}^{W\mu\text{CR}} + R_M \cdot n_{\text{fake}-\mu,i} \right) \\
\times \prod_i P \left( N_i^{W\nu\text{CR}} \mid \beta_i \cdot B_{W,i}^{W\nu\text{CR}} + B_{\text{non}-W,i}^{W\nu\text{CR}} + R_S,i \cdot n_{\text{fake}-\mu,i} \right),
\]

where \( \beta_i \) are the same free parameters present in the signal region yield likelihood to propagate the control region constraints to the expected \( W \) and \( Z \) background yields in the SR. The \( B_{W,i}^{W\mu\text{CR}} \) and \( B_{W,i}^{W\nu\text{CR}} \) terms are the simulated \( W+\text{jets} \) yields separated by lepton flavour, and \( B_{Z,i}^{Z\text{CR}} \) are the \( Z+\text{jets} \) yields in simulation scaled by the same theoretical model for the ratio of \( Z/W \) as in the SR. The \( B_{\text{non}-Z,i}^{Z\text{CR}}, B_{\text{non}-W,i}^{W\mu\text{CR}}, \) and \( B_{\text{non}-W,i}^{W\nu\text{CR}} \) are the simulated yields due to non-\( W \), non-\( Z \), and non-fake contributions to these control regions. The ‘fake’ electrons (muons) are modelled with free parameters \( n_{\text{fake}-\mu,i} \) (\( n_{\text{fake}-\mu,i} \)) multiplied by the \( R_S,i \) (\( R_M \)) scale factors introduced in section 5.1.

The last set of yield probabilities in the likelihood are for the fake-\( e \) and fake-\( \mu \) control regions

\[
\mathcal{L}^{\text{fake-CR}} = \prod_i P \left( N_i^{\text{fake}-\mu\text{CR}} \mid \beta_i \cdot B_{W,i}^{\text{fake}-\mu\text{CR}} + B_{\text{non}-W,i}^{\text{fake}-\mu\text{CR}} + n_{\text{fake}-\mu,i} \right) \\
\times \prod_i P \left( N_i^{\text{fake}-\mu\text{CR}} \mid \beta_i \cdot B_{W,i}^{\text{fake}-\mu\text{CR}} + B_{\text{non}-W,i}^{\text{fake}-\mu\text{CR}} + n_{\text{fake}-\mu,i} \right),
\]

where the first two terms in each correspond to the real-lepton contributions calculated in the same way as in the \( W \) and \( Z \) control regions, and the last terms, \( n_{\text{fake}-\mu,i} \) and \( n_{\text{fake}-\nu,i} \), are the contributions from the fake leptons.

To constrain the nuisance parameters within their uncertainties the following likelihood term is considered:

\[
\mathcal{L}^{NP} = \prod_j G(0|\theta_j),
\]

where \( G \) is the Gaussian probability density function. All systematic uncertainties are implemented as nuisance parameters \( \bar{\theta} \) modifying the \( B \), \( S \), and \( R \) parameters. Their nominal values are zero and they are constrained by Gaussian probability distributions with widths corresponding to the size of their uncertainties, which are normalised to one. The effect of the nuisance parameters on the yields is through an exponential response function \((1 + \varepsilon_{ij})^{\theta_j} \) where \( \varepsilon_{ij} \) is the fractional uncertainty amplitude of bin \( i \) from the uncertainty source \( j \). This treatment prevents the fitted yields from becoming negative in cases with large uncertainties.

Each experimental uncertainty source is taken to be fully correlated across all signal and control regions. For the correlation of theory systematic uncertainties the following assumptions are made. PDF uncertainties are treated as fully correlated across bins. The perturbative uncertainties of each of the four \( V+\text{jets} \) samples (strong and electroweak
components of $W$ and $Z$ production) are split into 17 components: one for each of the 16 SR (CR) bins, and one component that is fully correlated across all bins. Each of the 16 uncorrelated uncertainty sources only affects one SR (CR) bin, with an uncertainty amplitude obtained from the seven-point scale variation as discussed in section 6.1.1. The last uncertainty component, which affects all bins for a particular $V+$ jets sample, is assigned such that the combined uncertainty is equal to the systematic uncertainty for the inclusive SR (CR).

Combining all the terms gives the full likelihood expression

$$L(\mu, \vec{\beta}_V, \vec{n}_{PU-MJ/fake-\mu}, \vec{\theta}) = L^{SR} \cdot L^{PU-CR} \cdot L^{V+jets-CR} \cdot L^{fake-CR} \cdot L^{NP},$$

which contains a total of 58 free parameters and 105 observed event counts. Asymptotic formulae for the CL$_s$ frequentist approach [140–142] are used to set an upper limit on $\mu$ at 95% CL. This corresponds to the observed upper limit on the branching fraction of Higgs boson decays into invisible particles.

8 Validation

Since the event selection requires $\Delta \phi_{jj} < 2$, events with $2 < \Delta \phi_{jj} < 2.5$ are not part of the analysis but are instead used to define a validation region (VR). The VR data set is used to validate the background estimation techniques, and in particular to test the used ratio of the $W+$ jets and $Z+$ jets cross sections and the assigned uncertainties (see section 6.1.1). The validation parallels the full analysis chain with five VR bins in $m_{jj}$ with the same boundaries as the SR bins 1–5 or 6–10. The expected signal from invisibly decaying Higgs bosons at an assumed branching fraction $B_{inv}$ of 0.15 is less than 5% of the background in all VR bins corresponding to the SR. Because of the larger multijet background for bins with $E_T^{miss} < 200$ GeV or $N_{jets} = 3$ or 4, only events with $E_T^{miss} > 200$ GeV with $N_{jets} = 2$ are considered. Corresponding validation regions, each with five bins, are defined for $W+$ jets, $Z+$ jets, fake-$e$, and fake-$\mu$. The multijet background is estimated with the R+S method. The likelihood function is adjusted to the reduced number of VR bins. A fit is performed in the VR bins, excluding the zero-lepton bins, using only the background contributions and with $\mu$ set to zero. In total, the likelihood function for the VR has 15 free parameters and 30 observed event counts. The fitted $V+$ jets background scaling factors $\beta_i$ are then applied to the corresponding VR bins with zero leptons, and the resulting background expectations are compared with the observed yields as shown in figure 10. Good agreement between the total background prediction and data yields is obtained across all bins, including the zero-lepton validation region, which justifies the application of this strategy to the SR bins with $\Delta \phi_{jj} < 2$.

9 Results

The fit strategy as described in section 7 is applied to the data set. Overall, good agreement between the predicted and observed yields is observed. The expected background yields in the inclusive signal region and the control regions prior to the likelihood fit are summarised in
Figure 10. Post-fit yields and their uncertainties for all VR bins with $2 < \Delta \phi_{jj} < 2.5$ and with $\mu$ set to zero. The five bins in each region correspond to the $m_{jj}$ bins of the SR bins 1–5 and 6–10. The fit is performed only in the VR bins with one or two leptons, and the fitted background scaling factors $\beta_i$ are applied to the corresponding VR bins with zero leptons. Minor contributions from $t\bar{t}$, $VV$, $VVV$, and VBF $H \rightarrow \tau^+\tau^-$ or $H \rightarrow W^+W^-$ are combined and labelled ‘other’. The resulting background expectations are compared with the observed yield. The signal is scaled to a branching ratio of 15% for Higgs boson decays into invisible particles. The ratio of the observed yields to the prediction of the model post-fit is shown by the points in the bottom plot and the relative change of the prediction with respect to the pre-fit yields is shown by the blue line.

table 5. The event yields in the 16 signal selection bins after a background-only likelihood fit, i.e. with the signal strength $\mu$ set to zero, are compared with the observed numbers of events in tables 6 and 7, and a comparison in CR and SR bins is presented in figure 11, excluding the fake-lepton and pile-up CR bins. Figure 12 shows the $m_{jj}$ and $\Delta \phi_{jj}$ distributions after the likelihood fit in the inclusive signal region. The ratio of yields before and after the likelihood fit shows a clear trend in $m_{jj}$. This is in agreement with the observation that SHERPA 2.2.1 tends to overestimate the strong $Z + \text{jets}$ processes at large values of $m_{jj}$ [143].

The fitted value of the signal strength $\mu$ is $0.053 \pm 0.052$. The contributions of the systematic uncertainties to the 68% confidence interval of the fitted $\mu$ are shown in table 8. They are evaluated in a fit with the data yields set to the total background expectation, as is done to determine the expected limits. Groups of systematic uncertainties (section 6) are subsequently switched off, i.e. the corresponding nuisance parameters are fixed to their best-fit values. The reduction of the 68% confidence interval of $\mu$ is obtained by quadratically subtracting the resulting size of the confidence interval from the nominal case, which includes all systematic uncertainties. The statistical uncertainties of the data yields in the $V + \text{jets}$ CR bins and the remaining SR and CR bins have an impact of $\sim 42\%$. 
<table>
<thead>
<tr>
<th>Process</th>
<th>SR</th>
<th>$Z_{\ell\ell}$ CR</th>
<th>$W_\ell\nu$ CR</th>
<th>$W_\mu\nu$ CR</th>
<th>Fake-(e) CR</th>
<th>Fake-(\mu) CR</th>
<th>Pile-up CR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z$ strong</td>
<td>6030 ± 2050</td>
<td>1220 ± 440</td>
<td>42 ± 14</td>
<td>143 ± 21</td>
<td>160 ± 57</td>
<td>40 ± 14</td>
<td>42 ± 27</td>
</tr>
<tr>
<td>$Z$ EWK</td>
<td>2630 ± 260</td>
<td>618 ± 74</td>
<td>12.1 ± 1.8</td>
<td>28.3 ± 3.3</td>
<td>23.5 ± 3.3</td>
<td>15.3 ± 2.8</td>
<td>7.2 ± 3.1</td>
</tr>
<tr>
<td>$W$ strong</td>
<td>3710 ± 1300</td>
<td>-</td>
<td>3260 ± 1180</td>
<td>5710 ± 1850</td>
<td>1810 ± 650</td>
<td>1010 ± 370</td>
<td>49 ± 27</td>
</tr>
<tr>
<td>$W$ EWK</td>
<td>1610 ± 150</td>
<td>-</td>
<td>2360 ± 190</td>
<td>3410 ± 240</td>
<td>1400 ± 140</td>
<td>822 ± 69</td>
<td>22.9 ± 7.7</td>
</tr>
<tr>
<td>Fake-(e)</td>
<td>-</td>
<td>-</td>
<td>191 ± 70</td>
<td>-</td>
<td>1100 ± 330</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Fake-(\mu)</td>
<td>-</td>
<td>-</td>
<td>43 ± 15</td>
<td>-</td>
<td>130 ± 51</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Multijet</td>
<td>830 ± 190</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1890 ± 110</td>
<td>-</td>
</tr>
<tr>
<td>Other</td>
<td>180 ± 46</td>
<td>46 ± 25</td>
<td>346 ± 82</td>
<td>351 ± 71</td>
<td>67 ± 13</td>
<td>89 ± 27</td>
<td>19.5 ± 9.0</td>
</tr>
<tr>
<td>Total bkg.</td>
<td>14990 ± 2990</td>
<td>1880 ± 510</td>
<td>6210 ± 1260</td>
<td>9150 ± 1890</td>
<td>4560 ± 760</td>
<td>2110 ± 390</td>
<td>2630 ± 110</td>
</tr>
</tbody>
</table>

| $H$ (VBF)        | 886 ± 81    | 3.9 ± 1.3         |               |              |              |              |            |
| $H$ (ggF)        | 106 ± 41    | Predicted signal for $B_{\text{inv}} = 15\%$ | 1.0 ± 1.3     |              |              |              |            |
| $H$ (VH)         | 0.9 ± 0.2   |                    |                |              |              |              |            |

| Data             | 16 490      | 2051              | 6361           | 9294         | 4563         | 2110         | 2033       |

**Table 5.** Observed and expected background event yields with associated uncertainties in the signal region (SR) and control regions ($Z_{\ell\ell}$ CR, $W_\ell\nu$ CR, $W_\mu\nu$ CR, fake-\(e\) CR, fake-\(\mu\) CR, pile-up CR) prior to the likelihood fit. Minor backgrounds from $t\bar{t}$, $VV$, $VVV$, and VBF $H \rightarrow W^+W^-\tau^+\tau^-$ are combined and labelled ‘other’. The uncertainties in the backgrounds include the statistical, experimental, and theoretical uncertainties, taking into account the correlations between the individual SR and CR bins. The predicted signal yields (VBF, ggF, and $VH$) for $B_{\text{inv}} = 15\%$ (the observed limit) are presented for comparison. For all CRs with leptons, the contribution from the predicted signal is negligible.

and $\sim$54\%, respectively. For the experimental uncertainties, the dominant two groups are the uncertainties associated with the leptons (reconstruction and isolation efficiencies, lepton energy scale and resolution) and the uncertainties in the JER, having an impact of $\sim$32\% and $\sim$29\%, respectively. The uncertainties in the multijet background (section 5.2) contribute with $\sim$40\% and the contributions of fake leptons to the $W$ CRs have an impact of $\sim$37\%. The dominant theoretical uncertainties are in the $V$+jets background estimates and amount to $\sim$28\%. Given, that the fit does not show any significant signal contribution, an upper limit of 0.145 is set on the branching fraction of Higgs boson decays into invisible particles at 95\% confidence level. This is in agreement with the expected value of $0.10^{+0.041}_{-0.028}$ as shown in table 9. In comparison with the observed (expected) limit on $B_{\text{inv}}$ of 0.37 (0.28) from the previous search [27], this analysis shows a significant improvement.

The limit on $B_{\text{inv}}$ is related to a limit on the spin-independent WIMP-nucleon cross section [11, 12, 144, 145]. This allows the result obtained here to be compared with limits from a selection of the most sensitive direct-detection experiments [146–148] in figure 13. The relation between $B_{\text{inv}}$ and $\sigma_{\text{WIMP-nucleon}}$ is presented in the effective field theory framework assuming that the new-physics scale is $O(1)$ TeV, well above the scale probed at the Higgs boson mass [12]. To enable comparisons with the 90\% CL direct-detection limits, the observed 90\% CL upper limit of 12.7\% on $B_{\text{inv}}$ from this analysis is used, the expected limit being 8.7\%. For the scalar WIMP interpretation, cross sections
Table 6. Yields of data, signal and major backgrounds in the signal regions for each bin after the likelihood fit with \( \mu \) set to zero. Minor backgrounds from \( t\bar{t}, VV, VVV, \) and \( \text{VBF} \) \( H \to W^*W/\tau^+\tau^- \) are combined and labelled ‘other’. The uncertainties in the backgrounds are derived by the fit and take into account the correlations of the background uncertainties. The predicted signal yields (VBF, ggF, and \( VH \)) for \( B_{\text{inv}} = 15\% \) (the observed limit) are presented for comparison, along with their associated uncertainties, assuming that the experimental uncertainties are fully correlated between the ggF and VBF contributions. The \( \beta_V \) are the background normalisation factors.
Adding a renormalisable mechanism for generating the vector-like WIMP masses could predict a new scalar particle that mixes with the Higgs boson. This adds at least two free parameters to the model, for example its mass $m_2$ and the mixing angle $\alpha$. The dependence of the exclusion limit for the vector-like WIMP hypothesis on the mass $m_2$ is shown in figure 13. The uncertainty band in the plot uses the latest computation of the nucleon form factors [152]. The overlay shows the complementarity in coverage by the direct-detection experiments and the searches at colliders, such as the presented analysis.

The results are further interpreted as a search for invisible decays of heavy scalar particles acting as mediators to dark matter. The considered masses range from 50 GeV to 2 TeV, and the upper limit on the product of cross section and branching ratio to invisible particles ($\sigma^{\text{VBF}} \cdot B_{\text{inv}}$) is shown in figure 14. The derived limits become stronger for heavier mediator masses due to an accumulation of the signal events at larger values of $m_{ij}$, where the background yields are smaller. The 95% CL upper limit on $\sigma^{\text{VBF}} \cdot B_{\text{inv}}$ is 1.0 pb at a mediator mass of 50 GeV and strengthens to 0.1 pb for a mediator mass of 2 TeV.
Figure 11. Post-fit results of all SR and CR bins with $\mu$ set to zero. Minor contributions from $t\bar{t}$, $VV$, $VVV$, and VBF $H$ with $H \rightarrow \tau^+\tau^-$ or $H \rightarrow W^*W$ are combined and labelled ‘other’. The signal (dashed, red line) is scaled to a branching ratio of 15% for Higgs boson decays into invisible particles. The 16 bins in the SR and each of the three CRs are defined in figure 2. The fake-lepton and pile-up CRs, which contain an additional 41 bins that also enter the likelihood fit, are not shown, as this improves the readability. The ratio of the observed yields to the prediction of the model post-fit is shown by the points in the bottom plot and the relative change of the prediction with respect to the pre-fit yields is shown by the blue line.
Figure 12. Post-fit (a) $m_{jj}$ and (b) $\Delta\phi_{jj}$ distributions in the inclusive signal region. Minor contributions from $t\bar{t}$, $VV$, $VVV$, and VBF $H$ with $H \rightarrow \tau^+\tau^-$ or $H \rightarrow W^*W$ are combined and labelled ‘other’. The signal (dashed red line) is scaled to a branching ratio of 15\% for Higgs boson decays into invisible particles. In addition to the data-to-background ratio, the lower panels show the ratio of the background expectation before and after the likelihood fit, the relative size of the multijet background, and the signal-to-background ratio. The latter two quantities are shifted by one to use the same $y$-axis scaling.
Table 8. The contributions to the 68% confidence interval of the fitted signal strength $\mu$ from different sources of systematic uncertainty. They are evaluated in a fit with the data yields set to the total background expectation. Due to residual correlations between categories, the sum in quadrature of the individual contributions only approximately equals the total uncertainty. The uncertainty due to the number of data events ('Data stats.') is obtained by fixing all systematic uncertainties to their best-fit values. The entry 'V+jets data stats.' is the statistical uncertainty originating from data yields for the $W$ and $Z$ normalisation. The experimental uncertainties and the uncertainty related to the size of MC simulated samples ('MC stats.') are treated as separate categories. The ‘Remaining’ category contains $E_T^{\text{miss}}$, luminosity, and pile-up uncertainties, and a diboson uncertainty in the $Z\ell\ell$ CR.

Table 9. Observed and expected limits on $B_{\text{inv}}$ for a Higgs boson with a mass of 125 GeV calculated at the 95% CL for a $139\,\text{fb}^{-1}$ data set. The $\pm\sigma$ variations of the expected limit are also shown.
Figure 13. Upper limits on the spin-independent WIMP-nucleon cross section using Higgs portal interpretations of $B_{\nu \nu}$ at 90% CL vs $m_{\text{WIMP}}$. For the vector-like WIMP hypothesis, the dependence on the mass $m_2$ of the new scalar particle, which is often predicted by renormalisable models, is shown for two different values taken from ref. [151]. For comparison with direct searches for DM, the plot shows results from refs. [146–148]. The neutrino floor for coherent elastic neutrino-nucleus scattering is taken from refs. [153, 154] and assumes germanium as the target over the whole WIMP mass range. The dependence on the choice of target nucleus is relatively small, given the large range of cross sections shown.
Figure 14. Upper limit on cross section times branching ratio to invisible particles for a scalar mediator as a function of its mass. For comparison the VBF cross section at NLO in QCD, i.e. without the electroweak corrections, for a particle with SM Higgs boson couplings, multiplied by a $B_{\text{inv}}$ value of 15%, is overlaid.

10 Conclusion

A search for Higgs boson decays into invisible particles is presented using 139 fb$^{-1}$ of proton-proton collision data with a centre-of-mass energy of $\sqrt{s} = 13$ TeV collected between 2015 and 2018 by the ATLAS detector at the LHC. The VBF production mode is targeted by requiring two energetic jets that are not back-to-back in the azimuthal plane and exhibit a large rapidity separation, in addition to large values of the two-jet invariant mass and missing transverse momentum. Events with three or four jets are also considered if the additional jets are compatible with the hypothesis that they originate from final-state radiation from one of the VBF jets. The dominant $Z + \text{jets}$ and $W + \text{jets}$ backgrounds are evaluated using common control regions, which is made possible due to the use of a dedicated theoretical calculation at next-to-leading order in the phase space that is relevant for this analysis. Two independent data-driven techniques are used to estimate the multijet background.

Assuming the Standard Model cross section for the Higgs boson, an upper limit of 0.145 is set on the branching fraction $B_{\text{inv}}$ into invisible particles at 95% CL, with an expected limit of 0.103. This is a significant improvement on the previously published result using 36 fb$^{-1}$ of Run-2 data and originates from various changes. The increased size of the data sample and simulated samples, together with refinements of the lepton identification and acceptance results in a relative improvement of the expected limit by $\sim$36%. The extension of the requirement that both leading jets are not pile-up tagged to forward jets with $|\eta| > 2.5$ improves the sensitivity by $\sim$17%. Considering events with three or four jets and events with lower values of missing transverse momentum, along with the finer binning of the signal region reduces the upper limit by $\sim$19%. The changes to the multijet and $V + \text{jets}$ background estimation lead to further improvements of $\sim$10% and $\sim$6%, respectively.
The result is interpreted using Higgs portal models to exclude regions in the parameter space of $\sigma_{\text{WIMP-nucleon}}$ and $m_{\text{WIMP}}$ for various WIMP models. The obtained results are further interpreted as a search for invisible decays of new scalar particles with masses of up to 2 TeV, resulting in an upper limit on $\sigma^{\text{VBF}} \cdot B_{\text{inv}}$ of 1.0 pb for a mediator mass of 50 GeV and of 0.1 pb for 2 TeV.

Acknowledgments

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently.

We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; ANID, Chile; CAS, MOST and NSFC, China; Minciencias, Colombia; MEYS CR, Czech Republic; DNR and DNSRC, Denmark; IN2P3-CNRS and CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF and MPG, Germany; GSRI, Greece; RGC and Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MEIN, Poland; FCT, Portugal; MNE/IFA, Romania; JINR; MES of Russia and NRC KI, Russian Federation; MESTD, Serbia; MIZŠ, Slovenia; DSI/NRF, South Africa; MICINN, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, CANARIE, Compute Canada and CRC, Canada; COST, ERC, ERDF, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex, Investissements d’Avenir Idex and ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF, Greece; BSF-NSF and GIF, Israel; Norwegian Financial Mechanism 2014–2021, Norway; NCN and NAWA, Poland; La Caixa Banking Foundation, CERCA Programme Generalitat de Catalunya and PROMETEO and GenT Programmes Generalitat Valenciana, Spain; Göran Gustafssons Stiftelse, Sweden; The Royal Society and Leverhulme Trust, United Kingdom.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (U.K.) and BNL (U.S.A.), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in ref. [155].

We thank Jonas Lindert, Marek Schönherr, and Stefano Augusto Pozzorini for providing the full next-to-leading-order corrections to the $V+\text{jets}$ cross sections in the vector-boson-fusion signature while being associated with ATLAS.

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited. SCOAP³ supports the goals of the International Year of Basic Sciences for Sustainable Development.
References


ATLAS collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, 2008 JINST 3 S08003 [SPIRE].


[93] C. Anastasiou et al., High precision determination of the gluon fusion Higgs boson cross-section at the LHC, JHEP 05 (2016) 058 [arXiv:1602.00695] [INSPIRE].


School of Physics and Astronomy, University of Birmingham, Birmingham; United Kingdom

Department for Physics and Technology, University of Bergen, Bergen; Norway

Institut für Physik, Humboldt Universität zu Berlin, Berlin; Germany

Institute of Physics, University of Belgrade, Belgrade; Serbia

Institut de Física d’Altes Energies (IFAE), Barcelona Institute of Science and Technology, Barcelona; Spain

Department of Physics, University of Adelaide, Adelaide; Australia

Department of Physics, University of Alberta, Edmonton AB; Canada

Department of Physics, Ankara University, Ankara; Turkey

LAPP, Univ. Savoie Mont Blanc, CNRS/IN2P3, Annecy; France

APC, Université Paris Cité, CNRS/IN2P3, Paris; France

Department of Physics, University of Alberta, Edmonton AB; Canada

Department of Physics, University of Adelaide, Adelaide; Australia

Department of Physics, Bogazici University, Istanbul; Turkey

Department of Physics, Oswiecim University, Oswiecim; Poland

School of Physics and Astronomy, University of Birmingham, Birmingham; United Kingdom

Department of Physics, Bogazici University, Istanbul; Turkey

School of Physics and Astronomy, University of Birmingham, Birmingham; United Kingdom

Institut de Física d’Altes Energies (IFAE), Barcelona Institute of Science and Technology, Barcelona; Spain

Facultad de Ciencias y Centro de Investigaciones, Universidad Antonio Nariño, Bogotá; Colombia

Departmento de Física, Universidad Nacional de Colombia, Bogotá; Colombia